التّوازن الدَّقيق: بين مُكافحة التّضليل الرقمي واحترام حُريّة التّعبير

تقرير بحثي للجنة النّطاق الواسع حول:
«حُريّة التّعبير والتّصدي للمعلومات المُضلّلة على الإنترنت»

أيلول/سبتمبر 2020





التوازن الدَّقيق:
بين مُكافحة التّضليل الرقمي
واحترام حُريّة التّعبير

تقرير بحثي للجنة النطاق الواسع حول: «حُريّة التّعبير والتّصدي للمعلومات المُضلّلة على الإنترنت»



فريق التّحرير:

كالينا بونتشيفا وجولي بوزيتي

المؤلفون المُساهمون:

كالينا بونتشيفا	جامعة شيفيلد، المملكة المتحدة
جولي بوزيتي	المركز الدولي للصحفيين (الولايات المتحدة)؛ مركز حُريّة الإعلام، جامعة شيفيلد (المملكة المتحدة)؛ معهد رويترز لدراسة الصحافة، جامعة أكسفورد، (المملكة المتحدة)
دينيس تيسو	وكالة فرانس برس، فرنسا
تريشا ماير	جامعة فريجي بروكسل، بلجيكا
سام غريغوري	منظمة شاهد WITNESS، الولايات المتحدة
كلارا هانوت	مختبر الاتحاد الأوروبي للمعلومات المُضلّلة، بلجيكا
ديانا ماينارد	جامعة شيفيلد، المملكة المتحدة

نُشر في عام 2020 من قِبل الاتحاد الدولي للاتصالات Place des Nations, CH-1211 Geneva 20, Switzerland
ومنظمة الأمم المتحدة للتربية والعلم والثقافة، (اليونسكو)
7 Place de Fontenoy, 75352 Paris 07 SP, France

ISBN 978-92-3-100403-2



تم ترجمة هذا التقرير الى اللغة العربية بدعم من وزارة الخارجية لدولة قطر.

قائمة المُحتويات

الأشكال

الجداول

لا تُعبر التسـميات المُسـتخدمة وطريقـة عـرض المـواد فـي هـذا التقريـر البحثـي عـن أي رأي مهمـا كان مـن جانـب لجنـة النطـاق الواسـع المعنيـة بالتنميـة المُسـتدامة بشـأن الوضـع القانونـي لأيّ بلـد أو إقليـم أو مدينـة أو منطقـة أو لسـلطات أيّ منهـا، أو بشـأن ترسـيم حدودهـا أو تخومهـا. وإنّ الأفـكار والآراء المُعـرب عنهـا فـي هـذا التقريـر البحثـي هـي أفـكار المؤلفيـن؛ ولا تعكـس بالضـرورة آراء لجنـة النطـاق الواسـع المعنيـة بالتنميـة المسـتدامة ولا تلزمهـا.

أعضاء الفريق العامل المعنِي بحُريّة التّعبير والتّصدي للمعلومات المُضلّلة:

1. حصّة الجابر، رئيس مجلس إدارة Eshailsat
2. بوكار با، SAMENA
3. مُعز شكشوك، مسؤول ارتباط، اليونسكو
4. بيوتر دموشوسكي- ليبسكي، المنظمة الأوروبية لسواتل الاتصالات EUTELSAT
5. رامين غولوزاده، جمهورية أذربيجان
6. كارلوس مانويل ياركي أوريبي، أمريكا موفيل America Movil
7. بيبان كيدرون، مؤسسة 5 رايتس 5 Rights Foundation
8. روبرت كيركباتريك، مُبادرة «النبض العالمي» التابعة للأمم المتحدة UN Global Pulse
9. يي تشيونغ (داتو) لي إيستيك ISTIC
10. أدريان لوفيت، مؤسسة الشبكة العالمية The Web Foundation
11. فيليب ميتزغر، مدير عام وزارة الاتصالات السويسرية
12. بول ميتشل، مايكروسوفت
13. اسبيرانزا نديجي، جامعة كينياتا
14. باتريك نيريشيما، مسؤول ارتباط، مدير عام هيئة تنظيم المرافق العامة في رواندا
15. جوانا روبنشتاين، المؤسسة العالمية للطفولة World Childhood Foundation
16. أكيم شتاينر، برنامج الأمم المتحدة الإنمائي (عن طريق مينرفا نوفيرو - بيليك)

فريق الخُبراء المسؤول عن الإشراف:

فابريسيو بينيفوتو (جامعة ميناس جيرايس الاتحادية)
ديفينا فراو - ميغز (جامعة السُّوربون الجديدة - باريس 3)
شيريان جورج (جامعة هونغ كونغ المعمدانية)
كلير واردل (مؤسسة فيرست درافت First Draft)
هيرمان واسرمان (جامعة كيب تاون)

شُكر وتقدير:

غاي بيرغر (اليونسكو)
غويلهيرم كانيلا دي سوزا غودوا (اليونسكو)
أوسكار كاستيانوس (اليونسكو)
جو هيروناكا (اليونسكو)
آنا بولومسكا (الاتحاد الدولي للاتصالات)

سيدريك واخولز (اليونسكو)
مارتن ويكندن (اليونسكو)
جوانا رايت (جامعة شيفيلد)
شانشان شو (اليونسكو)

ردود أفعال إضافية وردت من المنظمات الأعضاء في مجموعة عمل لجنة النطاق الواسع:

برنامج الأمم المتحدة الإنمائي (منيرفا نوفيرو بيليك، روبرت أوب)
مبادرة النبض العالمي التابعة للأمم المتحدة (تشيس إيرني)

التّمهيد



أودري أزولاي
المُدير العام لليونسكو



د. حصّة الجبار
رئيس Eshailsat والرئيس المشاركة لمجموعة العمل

إنّ التواصل عبر الإنترنت ضروري في عالم اليوم كما أثبتت الأشهر القليلة الماضية. حيث يلعب الإنترنت دوراً رئيسياً في تسهيل وتيسير النفاذ والقدرة على الوصول إلى عمليتي التعليم والتعلُّم، والثقافة والمعلومات النوعية. فهو يسمح لنا بالعمل من المنزل والبقاء على تواصل بالعائلة والأصدقاء. والآن أكثر من أي وقت مضى، أصبح التواصل عبر الإنترنت في صلب مجتمعاتنا المترابطة ببعضها البعض.

ويتيح هذا الترابط المتزايد فرصاً أكبر. فهو يمكّن الأفراد عبر تزويدهم بالمعلومات والمعارف التي تدعم بدورها التنمية والديمقراطية. كذلك، يسمح بتنويع اللُّغات، وإجراء الأعمال التجارية. كما ويشجّعنا على تقدير الثقافات المختلفة.

ومع ذلك، يمكن أن يكون للتواصل عبر الإنترنت جانب مظلم أيضاً. فالمعلومات الخاطئة والرسائل المُضلّلة تتزايد على المنصات الرقمية. ويجري استغلال تصميم الخوارزميات من خلال سُلوكيات وحملات منظمة، بحيث يُوصى تلقائياً بمُحتوى يضر بحقوق الإنسان والحريات الأساسية في شريط الأحداث الجارية (news feeds) ونتائج البحث. وهذا لا يؤثر على ثقتنا في المؤسسات العامة فحسب، بل يُهدّد السّلام والصحة العامة أيضاً.

وقد سلّطت جائحة فيروس كورونا الضوء على هذه المسألة بشكل كبير. فمع انتشار الفيروس في جميع أنحاء العالم، انتشر سيل من الشائعات والمعلومات المغلوطة و/أو غير الدقيقة. ويبيّن تقريرنا، على سبيل المثال، أنّ مقطع فيديو واحد من كل أربعة مقاطع فيديو شعبية تتناول الجائحة على منصة يوتيوب YouTube يحتوي على معلومات مغلوطة. وفي الوقت الذي يعمل فيه العلماء في جميع أنحاء العالم على تطوير لُقاح ضد هذا الفيروس، وجدت دراسة أخرى أنّ هناك أكثر من 1300 صفحة مناهضة لأخذ اللّقاح على فيس بوك تضمّ حوالي 100 مليون مُتابع. ومع ذلك، فإنّ المعلومات الدقيقة ضرورية لإنقاذ الأرواح، لاسيّما خلال الأزمات والجوائح الصحية. فعلى حدّ تعبير المؤرخ «يوفال نواه هراري» في المقابلة التي أُجريت معه في مجلة «يونسكو كوريير»، فإنّ أفضل طريقة للدفاع عن أنفسنا ضد مُسببات الأمراض؛ هي المعلومات لا العُزلة. ولهذا السبب، أصبح خلق التوازن بين حُريّة التّعبير والنضال من أجل النفاذ إلى المعلومات الموثوقة اليوم مُهمّاً أكثر من أيّ وقت مضى.

وعلينا التطرق إلى هذه المسألة الآن، كما ولابدّ لنا من التسلُّح بالأدوات الصحيحة والمُناسبة. وهذا ما يهدف التقرير إلى تحقيقه وإفساح المجال له؛ من خلال تحديد وتحليل ما لا يقل عن 11 طريقة لتعزيز المعلومات عالية الجودة. وتشمل مجموعة الأدوات الناتجة عن ذلك طائفةً واسعةً من الاستجابات تتراوح بين التدابير السياسية والتشريعية إلى الجهود التكنولوجية والمبادرات التثقيفية والمتعلقة بالدراية الإعلامية.

ولا شك في أنّ الصحفيين المحترفين هُم في صميم مجموعة الأدوات هذه؛ فمن خلال تحديد المشكلات والتحقيق فيها، يمكنهم تعقب الأكاذيب وكشف زيفها والتعامل معها، وفي الوقت نفسه

ضمان ألا يصبح النقاش المشروع ضحية هذا الصراع ضد الأكاذيب. وتدافع اليونسكو عن الدور الأساسي الذي يلعبه الصحفيُّون في مُجتمعاتنا. فمن خلال تشجيع النقاش العام؛ يُساعدون على تعزيز الوعي في صفوف المواطنين.

ولمعالجة هذه القضايا، وُضعت هذه الوثيقة الفريدة والشاملة تحت رعاية لجنة النطاق الواسع للتنمية المُستدامة، والتي يشترك في رئاستها فخامة الرئيس «بول كاغامي» و«كارلوس سليم». وهنا لابدّ من تقديم الشُّكر لمجموعة العمل المعنية بحُريّة التّعبير والتصدي للمعلومات المُضلّلة والتابعة للّجنة؛ على دعمها هذا البحث العالمي الذي يأتي في الوقت المناسب.

ولا شك في أن سيل المحتويات المغلوطة والمُضلّلة، في هذه الأوقات العصيبة التي نمر بها اليوم، هو بالضبط ما لا يحتاج إليه عالمنا. ومع ذلك، وكما يبيّن هذا التقرير، يمكننا من خلال العمل معاً أن ندافع عن المعلومات الموثوقة؛ عالية الجودة، والنهوض بحُريّة التّعبير في الوقت نفسه. وهذه الوثيقة مثال على «التعاون الرقمي» الذي دعا إليه الأمين العام للأمم المتحدة.

ولذلك فإننا نُشجّع المفوضين وغيرهم من أصحاب المصلحة على الاستفادة الكاملة من هذا التقرير. فمعاً يمكننا أن نضمن أن تحقق لجنة النطاق الواسع المعنية بالتنمية المستدامة كامل إمكاناتها.

المُلخّص التنفيذي

في حزيران/يونيو 2020، دعت أكثر من 130 دولةً عضواً في الأمم المتحدة ومراقباً معتمداً؛ جميع الدول إلى اتخاذ خُطوات لمُكافحة انتشار المعلومات المُضلّلة، خاصةً خلال جائحة فيروس كورونا (UN Africa Renewal, 2020). وأكدت كافة هذه الجهات أنّ على هذه الاستجابات أن تكون مبنيةً على:

- حُريّة التّعبير؛
  - حُريّة الصحافة
  - التشجيع على اعتماد أسمى الأخلاقيات والمعايير في الصحافة؛
  - حماية الصحفيين وغيرهم من العاملين في الإعلام.
- وأن تسعى إلى تعزيز:
  - الدراية الإعلامية والمعلوماتية (MIL)؛
  - ثقة الجمهور في العلوم، والحقائق، ووسائل الإعلام المستقلة، والمؤسسات الحكومية والدولية.

كما وتم الاعتراف بالحاجة إلى اتخاذ إجراءات لمُكافحة المعلومات المُضلّلة في لجنة النطاق الواسع للتنمية المُستدامة المُشتركة بين الاتحاد الدولي للاتصالات واليونسكو. وقد أنشأت اللّجنة مجموعة عمل معنية بحُريّة التّعبير والتصدي للمعلومات المُضلّلة؛ والتي كُلّفت بدورها بإجراء هذه الدراسة العالمية الشاملة في عام 2019. وقد تم إجراء البحث الذي تقوم عليه هذه الدراسة بين أيلول/سبتمبر 2019 وتموز/يوليو 2020 من قِبل فريق من الباحثين الدوليين من تخصُّصات متنوعة.

إنّ «التوازن الدقيق: بين مُكافحة التّضليل واحترام حُريّة التّعبير» يستخدم مصطلح «المعلومات المُضلّلة» أو «التّضليل» للإشارة إلى المُحتوى المغلوط والمُضلّل الذي قد تكون له عواقب ضارة محتملة، بغض النظر عن النوايا أو السلوكيات الكامنة وراء اختلاق ونشر هذه الرسائل. ولا ينصبّ التركيز على التعاريف فحسب، بل كذلك على كيفية استجابة الدول والشركات والمؤسسات والمنظمات في جميع أنحاء العالم لهذه الظاهرة على نطاق واسع. ويتضمن هذا التقرير تصنيفاً جديداً يشمل 11 استجابةً، يسمح بفهم أزمة المعلومات المُضلّلة على نطاق دولي، بما في ذلك خلال جائحة فيروس كورونا. كما ويوفّر أداةً من 23 خُطوةً تمّ تطويرها لتقييم هذه الاستجابات للمعلومات المُضلّلة، بما في ذلك تأثيرها على حُريّة التّعبير (انظر أدناه).

ويخلُص التقرير البحثي إلى أنه لا يمكن التصدي للمعلومات المُضلّلة من دون معالجة المخاوف المتعلقة بحُريّة التّعبير، وهذا ما يُفسّر لماذا ينبغي أن تعزز الإجراءات التي تسعى إلى التصدي للمعلومات المُضلّلة هذا الحق بدلاً من انتهاكه. كما ويؤكد على أنّ النفاذ إلى معلومات موثوقة يمكن الاعتماد عليها، على غِرار المعلومات التي تنتجها الصحافة المستقلة الناقدة، يساهم في التصدي للمعلومات المُضلّلة. بالإضافة إلى ذلك، أنتجت هذه الدراسة إطاراً يمثّل دورة حياة المعلومات المُضلّلة – من التحريض واختلاق المعلومات إلى طرق النشر، إلى الأثر على أرض الواقع. ويشمل الفئات التالية: 1. المُحرّضون 2. الوُكلاء 3. الرسائل 4. الوسائط 5. الجهات المُستهدفة/المفسرة – وهذا ما يعرف باختصار «IAMIT».

ويمكن طرح سلسلة من التساؤلات المتتالية في مختلف مراحل دورة الحياة في ما يخص الجهات الفاعلة:

1. المُحرّضون:

من هم المُحرّضون المباشرون وغير المباشرين والمستفيدون من المعلومات المُضلّلة؟ ما هي علاقتهم بالوكيل (الوكلاء) (المذكور أو المذكورين أدناه)؟ لماذا يتم نشر المعلومات المُضلّلة – ما هو الدافع السياسي أو المالي مثلاً، هل أن الهدف هو تعزيز موقع معين أم غيرة في غير موضعها؟ ما هي الدوافع الأيديولوجية؟ إلخ، وذلك يتضمن البحث حيث أمكن في وجود نية في الإضرار أو التّضليل.

2. الوكلاء:

من الذي يقوم من ناحية تشغيلية باختلاق ونشر المعلومات المُضلّلة؟ ويثير هذا السُّؤال قضايا تتعلّق بإسناد الفاعل (في ما يخص الهُويّة الأصيلة)، ونوع الفاعل («شخص مؤثر»، فرد، مسؤول، مجموعة، شركة، مؤسسة) ومستوى التنظيم والموارد المستخدمة، ومستوى التّشغيل الآلي. وهنا يظهر استخدام لبعض السلوكيات – مثل استخدام بعض التقنيات على غرار روبوتات الإنترنت bots وشبكات الدُّمَى sock puppet networks والهُويّات الزائفة.

3. الرسائل:

ما الذي يجري نشره؟ من الأمثلة على ذلك الادعاءات أو السرديات الكاذبة، والصور ومقاطع الفيديو المأخوذة خارج إطارها أو التي تم تعديلها لغايات الغش، والصور الزائفة الفجّة deep fakes، وما إلى ذلك. هل تشمل الاستجابات الفئات التي تتضمن معلومات مضللة (على سبيل المثال: المُحتوى السياسي/الانتخابي)؟ ما الذي يُشكّل رسائل قد تكون ضارة/ضارة بالفعل أو هي ذات ضرر وخيم؟ كيف يتم المزج بين المُحتوى المغلوط أو المضلل وبين أنواع أخرى من المُحتويات، مثل: المُحتوى الصادق والمُحتوى الذي يُعزز الكراهية أو المُحتوى الترفيهي أو الآراء الشخصية؟ كيف يتم استغلال عالم المجهول من خلال تكتيكات التّضليل؟ هل أنّ هذه الرسائل تسعى إلى صرف الانتباه أو زعزعة مصداقية المُحتوى الصادق و/أو الجهات الفاعلة المنخرطة في البحث عن الحقيقة (على غرار الصحفيين والباحثين)؟

4. الوسائط:

- ما هي المواقع/الخدمات عبر الإنترنت ووسائل الإعلام الإخبارية التي تنتشر عليها المعلومات المُضلّلة؟ إلى أي مدى تقفز هذه المعلومات عبر هذه الجهات الوسيطة، بحيث يتم نشرها أولاً على «الويب المظلم» وينتهي بها الأمر في وسائل الإعلام العادية؟
- كيف تنتشر هذه المعلومات؟ ما هو نموذج أعمال والخصائص الخوارزمية والسياسية للموقع أو التطبيق/الشبكة الوسيطة التي يتم استغلالها؟ هل تسعى الاستجابات إلى معالجة التحيّز الخوارزمي الذي قد يؤدي إلى تفضيل المعلومات المُضلّلة؟ كذلك، هل من دليل على وجود سلوك مُعد له (بما في ذلك السُّلوك المُفتعل) يقوم باستغلال نقاط الضعف، ليصوّر مُحتوىً معيناً على أنه يتمتع بشعبية كبرى (أو حتى بانتشار واسع) في حين أنه في الواقع قد يكون قد حقق هذا الانتشار من خلال التّلاعب المتعمد على الخوارزميات؟
- هل أنّ الجهات الوسيطة تتصرف بطرق تتسم بقدر كافٍ من المُساءلة والشفافية وتتخذ الإجراءات الضرورية والمتناسبة للحد من انتشار المعلومات المُضلّلة؟

5. الجهات المستهدفة /المفسرة

- من يتأثر بالمعلومات المُضلّلة؟ هل أنّ هذه المعلومات تستهدف الأفراد والصحفيين والعلماء والأنظمة (مثل: العمليات الانتخابية، والصحة العامة، والمعايير الدولية)، والمجتمعات المحلية، والمؤسسات (مثل: مراكز البحوث)؛ أو المنظمات (بما في ذلك وسائل الإعلام الإخبارية)؛

- ما هي الاستجابات التي يتبعها هؤلاء على الإنترنت و/أو الإجراءات التي يتخذونها في الواقع خارج إطار الإنترنت؟ ويشمل هذا السُّؤال استجابات مختلفة على غرار عدم اتخاذ أي إجراء أو مشاركة المعلومات مما يؤدي إلى تأكيد محتواها أو الإعجاب بالمعلومات أو مشاركة المعلومات المُضلّلة بهدف كشف زيفها. هل من تقارير إخبارية غير نقدية (قد تؤدي إلى تحويل دور الصحفي/المنظمة الإخبارية المتواطئة من جهة مستهدفة بالمعلومات إلى وكيل تضليل)؟

- ما هي الاستجابات التي تسعى إلى تحديد الرسائل التي تندرج تحت المعلومات المضللةً والتحقُّق من هُويّة المُحرّضين والوكلاء، فضلاً عن تحديد النوايا والأهداف؟

- ما هي الاستجابات الرامية إلى تقييد ومُكافحة وكلاء حملات التّضليل والمُحرّضين عليها؟

- ما هي الاستجابات الرامية إلى الحد من إنتاج وتوزيع المعلومات المُضلّلة والسلوكيات ذات الصلة، التي تُنفذها الوسائط والإعلام بوجه خاص؛

- ما هي الاستجابات التي تهدف إلى دعم الجهات المستهدفة/المفسرة لحملات التّضليل.

بعدئذٍ، يُحدّد البحث أحد عشر نوعاً من الاستجابات ويقيّمها وهي تندرج في فئات أربع:

1. **استجابات تهدف إلى تحديد المعلومات المُضلّلة** (تهدف إلى تحديد المعلومات المُضلّلة وكشف زيفها وإظهار حقيقتها)

 i. رصد المعلومات والتحقُّق من الوقائع

 ii. الاستجابات الاستقصائية

2. **الاستجابات التي تستهدف مُختلقي المعلومات المُضلّلة وناشريها من خلال تغيير البيئة التي تحكم وتُشكّل سلوكهم**

 iii. الاستجابات التشريعية وما قبل التشريعية والسياساتية

 vi. الحملات الوطنية والدولية لمُكافحة المعلومات المُضلّلة

 v. الاستجابات الانتخابية

3. **الاستجابات التي تستهدف آليات الإنتاج والتوزيع** (المتعلقة بسياسات وممارسات المؤسسات الوسيطة للمُحتوى)

iv. الاستجابات المتعلقة بتنظيم المُحتوى

iiv. الاستجابات التقنية والخوارزمية

iiiv. الاستجابات الهادفة إلى حجب الربح المالي

4. **الاستجابات الموجهة إلى الجماهير المستهدفة من حملات التّضليل** (أي الهادفة إلى دعم «الضحايا» المحتملين للمعلومات المُضلّلة)

xi. الاستجابات المعيارية والأخلاقية

x. الاستجابات التثقيفية

ix. استجابات التمكين وتصنيف المصداقية

ويتضح أنّ هذه الاستجابات للمعلومات المُضلّلة غالباً ما تكون مكمّلةً لبعضها البعض. ففي الكثير من الأحيان مثلاً، نجحت التحقيقات التي أجراها بعض الصحفيين في كشف معلومات مضللة موجودة على الإنترنت لم تتمكن من كشفها (أو لم تعترف بوجودها) عمليات الرصد والتحقُّق من الوقائع التي أجرتها شركات التواصل عبر الإنترنت. وبعبارة أخرى، فإن الإجراءات التي تتخذها الشركات لوقف تناقل المعلومات المُضلّلة تعتمد جزئياً على العمل الاستقصائي الذي تقوم بها جهات فاعلة أخرى. وبالمثل، حتى لو ساعدت بعض الجهود على خفض توريد المعلومات المُضلّلة وتناقلها، لا تزال هناك حاجة إلى تمكين الجهات المستهدفة من مواجهة المُحتوى الذي يصلها؛ وبالتالي المساعدة في منع إعادة نشر المعلومات بشكل واسع.

وتخلُص الدراسة أيضاً إلى أنّ هناك حالات قد يعمل فيها نوع من الاستجابات ضد نوع آخر. ومن الأمثلة على ذلك الإفراط في التركيز على الاستجابات التي تعتمد على الإجراءات من أعلى إلى أسفل على حساب الاستجابات التي تعتمد على التمكين من أسفل إلى أعلى. إضافةً إلى ذلك، هناك ظاهرة اصطياد الصحفيين في شبكات محددة لوكلاء المعلومات المُضلّلة من خلال تجريم نشر أو توزيع المعلومات الكاذبة (على سبيل المثال عن طريق قوانين «الأخبار الزائفة»). وهذا يضر مباشرةً بدور الصحافة المستقلة والنقدية باعتبارها مناهضةً للمعلومات المُضلّلة. ومن الأمثلة المُماثلة؛ عدم قيام شركات التواصل عبر الإنترنت بإزالة الهجمات المليئة بالمعلومات المُضلّلة التي تُشن ضد الصحفيين على أساس احترام «حُريّة التّعبير». وبهذه الطريقة، فإن الفهم الخاص لحُريّة التّعبير يقوّض حُريّة الصحافة وسلامة الإعلام؛ وبالتالي عمل الإعلام ضد المعلومات المُضلّلة. وتشير هذه الأمثلة إلى ضرورة مواءمة التدخُّلات المختلفة، بدلاً من عملها في اتجاهات مختلفة. وبناءً على ذلك، تدعو هذه الدراسة إلى التشاور والتعاون بين أصحاب المصلحة المختلفين في مُكافحة المعلومات المُضلّلة. وهذا يتماشى مع نموذج اليونسكو لعالمية الإنترنت، الذي يدعم مبدأ ترسيخ حوكمة أصحاب المصلحة في القضايا الرقمية.

وتعترف الدراسة كذلك بضرورة اتباع مُقاربة متعددة الأوجه –بما في ذلك معالجة العوامل الاجتماعية والاقتصادية التي تعزز التّضليل، من خلال إعادة بناء العقد الاجتماعي وثقة الجمهور في المؤسسات الديمقراطية، وتعزيز التماسك الاجتماعي، لا سيّما في المُجتمعات التي تشهد استقطاباً شديداً، ومن خلال التصدي لنماذج الأعمال التي تزدهر بفضل المعلومات المُضلّلة المدفوعة مثل الإعلانات التي تتخطى الحدود والمحتويات الاحتيالية المتنكرة برداء الأخبار المشروعة أو الآراء المبنية على الواقع.

وبالنسبة إلى جميع الذين يسعون إلى التدخل للتصدي للمعلومات المُضلّلة، تحث هذه الدراسة على أن تدرج كل جهة فاعلة الرصد والتقييم المنهجيين ضمن أنشطة الاستجابة التي تضطلع بها. وينبغي أن تشمل هذه التدابير تقييم الفعالية والأثر على الحق في حُريّة التّعبير والحصول على المعلومات، بما في ذلك الحق في الخصوصية.

وتؤكد النتائج أيضاً على الحاجة إلى تعزيز الشفافية والإفصاح الاستباقي في كافة الاستجابات الـ11 الساعية إلى التصدي للمعلومات المُضلّلة. ويتماشى ذلك مع روح الغاية 16.10 من أهداف التنمية المستدامة التي تدعو إلى نفاذ الجمهور إلى المعلومات والحريات الأساسية.

ومن بين التدابير الأخرى، يُشجع هذا التقرير البحثي مجتمع النطاق الواسع والجهات المانحة على مواصلة الاستثمار في التحقُّق المستقل من الوقائع، وفي الصحافة المهنية النقدية، وفي تطوير وسائط الإعلام، وفي الدراية الإعلامية والمعلوماتية، لاسيّما من خلال التدخلات التثقيفية التي تستهدف الأطفال والشباب وكبار السن والفئات المستضعفة. كما ويدعو البحث الجهات الفاعلة إلى تعزيز الحفاظ على الخصوصية والنفاذ العادل إلى البيانات الرئيسية من شركات التواصل عبر الإنترنت، للسماح بإجراء تحليل مستقل لظهور وانتشار وتأثير المعلومات المُضلّلة الموجودة على الإنترنت على المواطنين حول العالم، وخاصةً في سياق الانتخابات والصحة العامة والكوارث الطبيعية.

وبالإضافة إلى هذه المقترحات الشاملة، تتوجه الدراسة إلى مجموعات أصحاب المصلحة الرئيسية بمروحة من التوصيات المتعلقة بالإجراءات الواجب اتخاذها في كل حالة. ومن بين هذه التوصيات، تم تسليط الضوء على التوصيات التالية هنا:

توصيات للمنظمات الحكومية الدولية والمنظمات الدولية الأخرى، حسب الاقتضاء:

- تعزيز المساعدة التقنية المقدمة إلى الدول الأعضاء بناءً على طلبها من أجل المساعدة في وضع أُطُر وسياسات ناظمة للتصدي للمعلومات المُضلّلة، بما يتماشى مع المعايير الدولية لحُريّة التّعبير والخصوصية. ويمكن أن ينطوي ذلك على التشجيع على اعتماد أداة تقييم الاستجابات للمعلومات المُضلّلة المؤلفة من 23 خطوةً والتي وُضعت لغايات هذه الدراسة (انظر أدناه).

- وفي حالة اليونسكو بشكل خاص، وولايتها المتعلقة بحُريّة التّعبير؛ تكثيف العمل الجاري حول المعلومات المُضلّلة بالشراكة مع منظمات الأمم المتحدة الأخرى والجهات الفاعلة العاملة في هذا المجال.

توصيات للدُّول:

- رفض تداول المعلومات المُضلّلة بشكل فعّال، بما في ذلك الالتزام بعدم التلاعب بالرأي العام بشكل مباشر أو غير مباشر – على سبيل المثال: عن طريق «حملات التأثير» التي تتجها أطراف ثالثة كشركات العلاقات العامة العاملة في مجال نشر «الدعاية المظلمة».

- مُراجعة ومُواءمة استجاباتها للمعلومات المُضلّلة، باستخدام الأداة المؤلفة من 23 خطوةً لتقييم القوانين والسياسات التي ستوضع نتيجةً لهذه الدراسة، وذلك للتأكد من امتثالها للمعايير الدولية لحقوق الإنسان (لا سيما حُريّة التّعبير؛ بما في ذلك النفاذ إلى المعلومات والحق في الخصوصية) وفي الوقت نفسه لرصد وتقييم استجاباتها.

- تعزيز الشفافية والإفصاح الاستباقي عن المعلومات والبيانات الرسمية، ورصد هذا الأداء بما يتماشى مع الحق في المعلومات والمؤشر 16.10.2 من أهداف التنمية المستدامة الذي يقيّم اعتماد وتنفيذ الضمانات الدستورية والقانونية و/أو السياستية لحق نفاذ الجمهور إلى المعلومات.

توصيات للأحزاب السياسية وغيرها من الجهات السياسية الفاعلة:

- تسليط الضّوء على الخطر الذي يُشكّله اللاعبون السياسيون كمصادر وعوامل مضخمة للمعلومات المُضلّلة والعمل على تعزيز البيئة التي تحيط بالمعلومات والثقة بالمؤسسات الديمقراطية.

- الامتناع عن استخدام أساليب التّضليل في الحملات السياسية، بما في ذلك استخدام أدوات سرية للتلاعب بالرأي العام وشركات العلاقات العامة العاملة في مجال «الدعاية المظلمة/الهدّامة».

توصيات لهيئات مُراقبة الانتخابات والسُّلطات الوطنية:

- تحسين شفافية الإعلانات الانتخابية التي تعتمدها الأحزاب السياسية والمُرشّحون والمنظمات التابعة لهم من خلال اشتراط وضع قواعد بيانات عن الإعلانات؛ شاملة ومتاحة بشكل علني، والكشف عن إنفاق الأحزاب السياسية والمجموعات الداعمة لها.

- العمل مع الصحفيين والباحثين في عمليات التحقُّق من الوقائع والتحقيقات حول شبكات المعلومات المُضلّلة الانتخابية ومنتجي «الدعاية المظلمة».

توصيات لسُلطات إنفاذ القانون ولمؤسّسة القضاء:

- ضمان معرفة العاملين في إنفاذ القانون بالحق في حُريّة التّعبير والحق في احترام الخصوصية، بما في ذلك الحماية الممنوحة للصحفيين الذين ينشرون معلومات قابلة للتحقُّق؛ تهدف إلى خدمة المصلحة العامة، وتجنب اتخاذ إجراءات تعسفية في تطبيق القوانين التي تجرّم المعلومات المُضلّلة.

- بالنسبة إلى القُضاة وغيرهم من الجهات المعنية والتابعة لمرفق القضاء: التنبُّه عند مراجعة والنظر في القوانين والقضايا المتصلة بتدابير مُكافحة المعلومات المُضلّلة، مثل التجريم، وذلك لضمان الاحترام الكامل للمعايير الدولية المتعلقة بحُريّة التّعبير واحترام الخصوصية في إطار تلك التدابير.

توصيات لشركات التواصل عبر الإنترنت:

- العمل معاً في إطار حقوق الإنسان، للتصدي للمعلومات المُضلّلة العابرة للمنصّات، من أجل تحسين القدرات التكنولوجية في مجال الكشف عن المُحتوى الكاذب والمُضلّل؛ والحد منه بشكل أكثر فعالية، وتبادل البيانات حول هذا الموضوع.

- تطوير استجابات متعلقة بتنظيم المُحتوى تُمكّن المستخدمين من النفاذ بسهولة إلى الصحافة كمصدر للمعلومات القابلة للتحقُّق، والتي يتم نشرها لتحقيق المصلحة العامة، مع إعطاء الأولوية للوكالات الإخبارية التي تمارس الصحافة المُستقلة النقدية والتي تتمتّع بالأخلاقيات المهنيّة.

- الاعتراف بأنه إذا كان بالإمكان التصدي للمعلومات المُضلّلة والمغلوطة على وجه السُّرعة خلال جائحة ما؛ بسبب الخطر الجسيم الذي تُشكّله على الصحة العامة، فلابدّ أيضاً من اتخاذ إجراءات ضد التّضليل السياسي -خاصةً عندما يتقاطع مع خطاب الكراهية - حيث أنه قد يشكّل خطراً على الأرواح أيضاً. وهذا ما ينطبق كذلك على المعلومات المُضلّلة المتصلة بتغيُّر المناخ.

- الاعتراف بأهمية حُريّة الصحافة وسلامتها؛ كونهما مُكونين أساسيين للحق في التّعبير المُكرَّس والمنشود دُولياً، ممّا يعني أنه لا يمكن السماح بالعنف المُمارَس عبر الإنترنت ضد الصحفيين (وهو من السمات الرائجة لحملات التّضليل).

- تطبيق التحقُّق من الوقائع على كافة المحتويات السياسيّة (بما في ذلك الإعلانات، والرأي القائم على الحقائق، و«الحديث المباشر») التي ينشرها السياسيُّون والأحزاب السياسيّة والشّركات التابعة لها، والجهات الفاعلة السياسية الأخرى.

كما تتوجه الدراسة بتوصيات إلى أصحاب المصلحة الآخرين؛ على غرار وسائل الإعلام (بشقّيه المطبوع والإلكتروني)، ومنظمات المجتمع المدني، ووسائط الإعلان، والباحثين.

ويؤكد هذا التقرير البحثي عموماً أنّ مسائل حُريّة التّعبير والنفاذ إلى المعلومات والاعلام الناقد والمستقل - مع نفاذ مفتوح ومتاح بالمجّان على الإنترنت - ليست حقوقاً أساسيةً من حقوق الإنسان فحسب، بل أدوات أساسية في الترسانة المستخدمة لمُكافحة المعلومات المُضلّلة - سواءً أكانت هذه المعلومات مرتبطة بالوباء أو بالانتخابات أو بتغيّر المناخ أو بالقضايا الاجتماعية ذات الصّلة. وتسعى هذه الدراسة ذات التوقيت المُلائم إلى دعوة جميع أصحاب المصلحة لدعم هذه المعايير الدولية التي تتعرض - وأهداف التنمية المستدامة بصورةٍ أوسع - إلى تهديد كبير بفعل المعلومات المُضلّلة.

وهـي تحـذّر مـن أنّ مُكافحـة المعلومـات المُضلّلـة ليسـت دعـوةً لقمـع تعدُّديـة المعلومـات والـرأي، ولا لقمـع النقـاش السياسـتي النابـض، بل هـي كفـاح لتغليـب الحقيقـة؛ حيـث أنـه بغيـاب المعلومـات القائمـة علـى الأدلـة علـى نطـاق واسـع، لـن يكـون مـن الممكـن النفـاذ إلـى معلومـات موثوقـة وذات مصداقيـة وقابلـة للتحقـق بشـكل مسـتقل مـن شـأنها دعـم الديمقراطيـة والمسـاعدة علـى تفـادي تفاقـم آثـار الأزمـات علـى غـرار الأوبئـة.

ولا يجـب أن تقـوم «مُعالجـات» المعلومـات المُضلّلـة بمُفاقمـة «الـداء»، كمـا عليهـا ألا تخلـق تحديـات أسـوأ مـن التحـدي الرئيـس نفسـه. ومـن خـلال العمـل معـاً، يمكـن للجهـات الفاعلـة الضالعـة فـي تنفيـذ المبـادرات فـي إطـار الاسـتجابات الإحـدى عشـرة المشـمولة فـي هـذه الدراسـة أن تضمـن تمتُّـع إجراءاتهـا بالشـفافية ومراعاة الاعتبـارات الجنسـانية المرتبطـة بالنـوع الاجتماعـي، واحتـرام حقـوق الإنسـان. كمـا ويمكنهـا أن تعمـل علـى أن تكـون هـذه الإجـراءات خاضعـة للتقييـم بشـكل منهجـي وفعالـةً علـى النحـو الأمثـل.

إطار تقييم الاستجابات الهادفة إلى مواجهة المعلومات المُضلّلة

تُقـدّم الدراسـة إطـاراً لتقييـم حُريّـة التّعبيـر فـي الاسـتجابات الهادفـة إلـى مواجهـة المعلومـات المُضلّلـة، وذلـك لمسـاعدة الـدُّول الأعضـاء فـي اليونسـكو وغيرهـا مـن المؤسسـات علـى صياغـة اسـتجابات تشـريعية وتنظيميـة وسياسـية للتصـدي للمعلومـات المُضلّلـة بطريقـة تُعـزّز مـن حُريّـة التّعبيـر. وتشـمل الأداة 23 نقطـةً مرجعيـةً تسـمح بتقييـم الاسـتجابات وفقـاً للمعاييـر الدوليـة لحقـوق الإنسـان؛ مـع إيـلاء اهتمـام خـاص بالحـق فـي النفـاذ إلـى المعلومـات والحـق فـي احتـرام الخصوصيـة.

1. هل خضعت الاستجابات قبل الصياغة والتنفيذ لمشاركة وإبداء الرأي من أصحاب المصلحة المتعددين (خاصةً منظمات المجتمع المدني والباحثين المتخصصين وخبراء حُريّة الصحافة)؟ وفي ما يخص الاستجابات التشريعية، هل وُجدت فرصة مناسبة للتداول بها قبل اعتمادها، وهل يمكن إجراء استعراض مستقل لها؟

2. هل تحدد الاستجابات بوضوح وشفافية التحديات التي يتعيّن مُعالجتها (مثل: اللامبالاة الفردية أو أنشطة الاحتيال؛ عمل شركات التواصل عبر الإنترنت والمنظمات الإعلامية؛ ممارسات المسؤولين أو الجهات الفاعلة الأجنبية التي تؤثر سلباً على الصحة العامة وعلى سلامة الانتخابات وعلى التخفيف من حدّة آثار تغيّر المناخ.. إلخ)؟

3. هل تتضمّن الاستجابات تقييماً للأثر في ما يتعلق بالعواقب المترتبة على الأُطُر الدُّوليّة لحقوق الإنسان التي تعزز من حُريّة التّعبير أو حُريّة الصحافة أو النفاذ إلى المعلومات أو احترام الخصوصية؟

4. هل تؤثّر الاستجابات أو تحُد من الحق في حُريّة التّعبير والحق في احترام الخصوصية والحق في النفاذ إلى المعلومات؟ إذا كان الأمر كذلك، وتعتبر الظروف التي تؤدي إلى الاستجابة مناسبةً لمثل هذا التدخل (على سبيل المثال: جائحة فيروس كورونا)، فهل إنّ التدخل في هذه الحقوق هو معرّف بشكل ضيق وضروري ومتناسب ومحدد زمنياً؟

5. هل تقيّد استجابة معينة الأعمال الصحفية أو تعرّضها للخطر، ومنها على سبيل المثال: الإبلاغ والنشر والحفاظ على سرية المصادر؟ وهل تقيّد من الحق في النفاذ إلى المعلومات التي تصب في المصلحة العامة؟ وقد تشمل الاستجابات في هذه الفئة ما يلي: القوانين المتعلقة «بالأخبار الزائفة» والقيود المفروضة على حُريّة التنقل والنفاذ إلى المعلومات بشكل عام أو كما هو مطبق على موضوع معين (مثل: الإحصاءات الصحية والنفقات العامة)؛ واعتراض الاتصالات (أو التنصُّت خارج إطار القانون) والمراقبة الموجهة أو الجماعية؛ والاحتفاظ بالبيانات وتسليمها. وإذا كانت هذه التدابير تمس بالفعل بهذه الوظائف الصحفية أو بمُساءلة المكلفين بالمسؤولية أمام أصحاب الحقوق بوجهٍ عام، يُرجى الرُّجوع إلى النقطة الرابعة أعلاه.

6. إذا اتضح أنّ استجابةً معينةً تحُد من أيّ من الحقوق المبينة في النُّقطة الرابعة، هل توفر استثناءات للأعمال الصحفية؟

7. هل تُؤخذ هذه الاستجابات بالحسبان (على سبيل المثال: التثقيفية، المعيارية، القانونية.. إلخ) مُجتمعةً وبشكل كلي من حيث أدوارها المختلفة وأوجه تكاملها والتناقُضات المُحتملة؟

8. هل لهذه الاستجابات طابع مقيّد بشكل حصري (على سبيل المثال: القيود القانونية على حملات التّضليل الانتخابية)، أم أنها تُقيم توازناً مناسباً بين التدابير المقيّدة والتدابير التمكينيّة (مثل: تثقيف الناخبين وتعزيز الدراية الإعلامية والمعلوماتيّة)؟

9. في حين قد تكون المعلومات المُضلّلة والمعلومات المغلوطة بالدرجة نفسها من الخطورة من حيث الأثر، هل تعترف الاستجابات بالفرق في الدوافع بين الجهات المتورطة في الخطأ المُتعمَّد (المعلومات المُضلّلة) والجهات المتورطة في الخطأ غير المُتعمَّد (المعلومات المغلوطة أو غير الدقيقة)، وهل تُصمَّم الإجراءات وفقاً لذلك؟

10. هل تخلط الاستجابات بين مُحتوى المعلومات المُضلّلة وبين مُحتوى خطاب الكراهية أو تُساوي بينهما (على الرغم من أنّ المعايير الدُّوليّة تُبرّر تدخُّلات قويّة للحد من خطاب الكراهية، في حين أن المعلومات المغلوطة ليست مُستبعدةً بحد ذاتها من حُريّة التّعبير)؟

11. هل يتمتّع الصحفيُّون والجهات الفاعلة السياسية والمُدافعون عن حقوق الإنسان بالحماية القضائية الفعالة من حملات التّضليل و/أو الكراهية التي تُحرّض على العداء والعنف والتمييز، والتي تهدف إلى ترهيبهم؟

12. هل تترافق الاستجابات القانونية بتوفير توجيه وتدريب على التنفيذ لسُلطات إنفاذ القانون والمُدَّعين العامين والقضاة في ما يتعلق بحماية الحق في حُريّة التّعبير، وهو حق أساسي، والآثار المترتبة على تقييد هذا الحق؟

13. هل يمكن تقييم الاستجابة بشفافية، وهل هناك عملية لرصد وتقييم آثارها على حُريّة التّعبير بشكل منهجي؟

14. هل الاستجابات هي موضوع تدابير رقابة ومُساءلة، بما في ذلك أنظمة المُراجعة والمساءلة (على غِرار تقديم التقارير إلى الجمهور والبرلمانيين وأصحاب المصلحة المعنيين)؟

15. هل يمكن الطعن في استجابة معينة أو سحبها إذا تبيّن أنّ آثارها السلبية على الحق في حُريّة التّعبير والحق في النفاذ إلى المعلومات والحق في الخصوصية تفوق فوائدها (علماً أنّ هذه الحقوق تُشكّل في حد ذاتها ترياقاً للتضليل)؟

16. هل وضعت التدابير المتعلقة بشركات ومنصات التواصل عبر الإنترنت مع إيلاء الاعتبار اللازم لمشاركة أصحاب المصلحة ولتعزيز الشفافية والمُساءلة، مع تجنُّب خصخصة الرقابة في الوقت عينه؟

17. هل هناك تقييم (يستعين بمشورة الخبراء) لإمكانات وحدود الاستجابات التكنولوجية التي تتناول المعلومات المُضلّلة (مع عدم المساس بحُريّة التّعبير والخصوصية)؟ هل هناك توقُّعات غير واقعيّة في ما يخص دور التكنولوجيا؟

18. هل تُشارك الجهات الفاعلة من المُجتمع المدني (بما في ذلك المنظمات غير الحكوميّة والباحثون ووسائط الإعلام) كشريك مستقل في ما يخص مُكافحة المعلومات المُضلّلة؟

19. هل تدعم الاستجابات إنتاج المعلومات وتوريدها وتداولها – بما في ذلك المعلومات المحلية ومتعددة اللُّغات – كبديل موثوق به للمعلومات المُضلّلة؟ ومن الأمثلة على ذلك: تقديم إعانات للصحافة الاستقصائية العاملة على التحقيق في المعلومات المُضلّلة، ودعم الإذاعات في المجتمعات المحلية ووسائط الإعلام الناطقة بلغات الأقليات.

20. هل تشمل الاستجابات تقديم الدعم للمؤسسات (على غِرار إعلانات ورسائل التوعية العامة في الإعلام؛ والمدارس) من أجل تعزيز العمل الهادف إلى التصدي للتضليل؟ قد يشمل ذلك تدخلات كالاستثمار في المشاريع والبرامج المُصمَّمة خصيصاً للمُساعدة على «تحصين» المجتمعات المحلية الكبيرة ضد المعلومات المُضلّلة من خلال برامج الدراية الإعلاميّة والمعلوماتيّة.

21. هل تُعظّم الاستجابات من انفتاح وتوافُر البيانات التي تحتفظ بها السُّلطات الحكومية، مع مراعاة واجب حماية واحترام الخصوصية الشخصيّة، كجزء من الحق في النفاذ إلى المعلومات والعمل الرسمي الهادف إلى استباق الشائعات وتمكين البحوث وتقديم التقارير المبنية على الحقائق؟

22. هل تُراعي الاستجابات الاعتبارات الجنسانية المرتبطة بالنوع الاجتماعي وأوجه الضعف الخاصة في بعض المجموعات (مثل: الشباب والمسنين) التي تجعلها أكثر هشاشةً في ما يخص التعرُّض للمعلومات المُضلّلة ونشرها وآثارها؟

23. إذا اتخذت التدابير الخاصة لمواجهة تحدٍّ مُلحّ، أو إذا كانت مُصمَّمةً ليكون لها أثر على المدى القصير (مثل: التدخّلات المقيّدة في الزمن المرتبطة بالانتخابات) فهل تصحبها مبادرات أو برامج أو حملات تهدف إلى إحداث التغيير وترسيخه على المديين المتوسط والطويل؟

1

الفصل الأول: المُقدّمة

المُؤلّفون: كالينا بونتشيفا، جولي بوزيتي



تسعى هذه الدراسة العالمية إلى مسح مختلف الاستجابات الدولية للمعلومات المُضلّلة وتعزيز فهمها، وإلى مسح وفهم آثار التدابير المناهضة للمعلومات المُضلّلة على الحق في حُريّة الرأي والتّعبير، المعرّف في المادة 19 من إعلان الأمم المتحدة العالمي لحقوق الإنسان[1]:

> **لكل فرد الحق في حُريّة الرأي والتّعبير؛ ويشمل هذا الحق في حُريّة اعتناق الآراء من دون تدخل، والتماس المعلومات والأفكار وتلقيها ونقلها من خلال أي وسيلة من وسائل الإعلام، دونما اعتبار للحدود.**

إنّ الحق في حُريّة التّعبير، بما في ذلك حُريّة الصحافة والحق في النفاذ إلى المعلومات، مكرّس كما الحق في احترام الخصوصية وذلك في القانون الدولي لحقوق الإنسان. لذلك، عند الاقتضاء، تتناول هذه الدراسة أيضاً قضايا احترام الخصوصية وصون الكرامة على شبكة الإنترنت. إضافةً إلى ذلك، فإنها تضع أزمة المعلومات المُضلّلة في سياق الدور التمكيني الذي يلعبه الإنترنت –خاصةً شبكات التواصل الاجتماعي – لفائدة تعزيز النفاذ إلى المعلومات ونقل المعلومات المُضلّلة في الوقت نفسه. وتُناقش الدراسة بالتفصيل إمكانية قيام الاستجابات المتصدية للمعلومات المُضلّلة بكبح حُريّة التّعبير، وتقترح سُبلاً لتجنب مثل هذه الآثار.

وعلى الرغم من أنّ العديد من الدراسات والأوراق السياستية المتعلقة بالمعلومات المُضلّلة قد تم نشرها من قِبل حكومات ومنظمات دولية وأكاديميين ومراكز فكر مستقلة، فإنّ هذه الدراسة تقدم إسهامات جديدة من خلال تطويرها لتصنيف منهجي لمُجمل الاستجابات للمعلومات المُضلّلة المعتمدة دولياً:

1. هي تعالج مجموعة الاستجابات الكاملة، بدلاً من الاستجابات التثقيفية أو القانونية أو التكنولوجية فقط على سبيل المثال؛
2. كما وتُصنّف الاستجابات وفقاً لهدف التدخل، وليس من حيث الوسائل المستخدمة أو الجهات الفاعلة المعنية؛
3. وتقيّم الاستجابات من حيث الافتراضات الأساسية والمعنى من منظار حُريّة التّعبير؛
4. تمثل قضايا وحالات واستجابات متنوعة جغرافياً، بما في ذلك التركيز على دول الجنوب؛
5. تقدم لمحةً عامةً عن الاستجابات الهادفة إلى «تسطيح المنحنى» لأزمة المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا (Posetti & Bontcheva 2020a; Posetti & Bontcheva 2020b)

هناك تعريفات متنوعة للمعلومات الخاطئة والمُضلّلة، ولكن لأغراض هذه الدراسة يستخدم مصطلح «المعلومات المُضلّلة» للإشارة بشكل عام إلى المُحتوى غير الصحيح الذي قد تترتب عنه آثار ضارة محتملة، على سبيل المثال، على صحة وسلامة الأفراد وعلى حُسن سير الديمقراطية. وبالنسبة إلى العديد من المُحلّلين، فإنّ نية الوكيل الذي يختلق المُحتوى غير الدقيق أو يُشاركه قد يساعد أيضاً على التمييز بين المعلومات المُضلّلة (الخطأ المتعمد) والمعلومات المغلوطة (الخطأ غير المتعمد). وتقبل هذه الدراسة بدور

1 https://www.un.org/en/universal-declaration-human-rights/

هذا التمييز، الذي ينطوي أيضاً على أنواع مختلفة من المعالجات. ومع ذلك، قد يكون أثر المُحتوى غير الصحيح والمُحتوى المضلل هو نفسه، بغض النظر عن النوايا التي تقف وراءه. ولا شكّ في أن هذا التركيز على الآثار الضارة المحتملة للمُحتوى غير الصحيح والمضلل لا على الدافع وراء اختلاقه ونشره؛ هو ما يفسر استخدام الدراسة لمصطلح المعلومات المُضلّلة كمصطلح جامع – بغض النظر عن القصد أو السُّلوك الكامن وراء نشر هذه الرسائل. ويرد مزيد من الشرح لهذه المسألة في القسم 1.2 المتعلق بالتعاريف أدناه.

وقد تسبب المعلومات المُضلّلة (على عكس المعلومات التي يمكن التحقُّق منها) ضرراً؛ لأنها قد تؤدي إلى إرباك المواطنين أو التلاعب بهم كما وإلى زعزعة الثقة في المعايير الدولية والمؤسسات أو في الاستراتيجيات المتفق عليها ديمقراطياً. كما وأنها قد تؤدي إلى تعطيل الانتخابات أو رسم صورة زائفة للتحديات الأساسية كتغيُّر المناخ. كما ويمكن أن تكون هذه المعلومات مُميتة، كما يتجلى ذلك من أزمة المعلومات المُضلّلة بخصوص جائحة فيروس كورونا (Posetti and Bontcheva 2020a, Posetti and Bontcheva 2020b).

وعـادةً مـا يتـم نشـر المعلومـات المُضلّـلة علـى يـد جهـات حكوميـة وغيـر حكوميـة، بمـا فـي ذلـك الأفـراد والجماعـات المنظمـة. فيتـم اختلاقها ونشـرها وتضخيمها بطريقـة طبيعيـة، علـى يد أفـراد يؤمنون بصحتها، أو بطريقـة مصطنعـة من خـلال الحمـلات التي تسـتخدم الأدوات التكنولوجية علـى غـرار روبوتـات الإنترنـت bots والخوارزميـات. وقـد وُضعـت بطريقـة تسـمح باسـتغلال التحيـزات المعرفيـة كالتحيـزات المبنيـة علـى لفـت الانتبـاه والتحيـزات التأكيديـة، وهـي تسـتخدم تقنيـات الدعايـة الشـعبية الزائفـة المعروفـة بعبـارة «astroturfing»[2] لتحفيـز مـا يُعـرف بأثـر bandwagon effect «سـلوك القطيـع» (Schmitt-Beck, 2008) وذلـك لخلق الانطبـاع أنّ هنـاك إيمـان واسـع بصحـة مسـألة أو غـرض معيـن. وغالبـاً مـا تسـعى حمـلات التّضليل إلـى اسـتهداف مـن أنتـج معلومـات قابلـة للتحقق أو مـن لديـه آراء معارضـة والتشـكيك فـي مصداقيته، بمـا يتضمّـن السياسـيين والصحفييـن والناشـطين فـي مجـال حقـوق الإنسـان والعلمـاء وغيرهـم. وينفـذ العديـد مـن وكلاء المعلومـات المُضلّـلة حمـلات مُعد لهـا عابـرةً للمنصـات، تصحبهـا تهديـدات وأسـاليب ترهيبيـة وتخريبيـة.

وعلـى وجـه الخصـوص، تؤثـر المعلومـات المُضلّـلة تأثيراً سـلبياً علـى حقوق المواطنين في احتـرام الخصوصية وحُريّـة التّعبيـر والنفـاذ إلـى المعلومـات. ولكـنّ العديـد مـن الجهـود الراميـة إلـى التصـدي للمعلومـات المُضلّـلة عبـر الإنترنـت قـد تتعـارض أيضـاً مـع حقـوق الإنسـان الأساسـية هـذه، كمـا هـو مبيّـن فـي أقسـام هـذه الدراسـة. ولذلـك، لابـد لـلأدوات والتدابيـر والسياسـات الراميـة إلـى معالجـة أزمـة المعلومـات المُضلّـلة أن تضمـن حمايـة حقـوق المواطنيـن وتمثيـل مصالحهـم. وهـذا يعنـي اتبـاع مقاربـة تقـرّ بكيفيـة تأثير هـذه القضايـا علـى أصحـاب المصلحـة كالجهـات الفاعلـة الصحفيـة ومنظمـات المجتمـع المدنـي وشـركات التواصـل عبـر الإنترنـت[3]. ولكـن غالبـاً مـا ينشـأ توتـر بيـن هـذه الحقـوق والمصالـح مـن جهـة والجهـود الهادفـة إلـى تحديـد

[2] ‘Astroturfing’ هو مصطلح مأخوذ عن اسم علامة تجارية للعُشب الاصطناعي المستخدم لتغطية الأسطح الخارجية؛ بما يجلب الانطباع بأنه غطاء عُشب طبيعي. أما في سياق المعلومات المُضلّلة؛ ينطوي هذا المصطلح على وضع بذور المعلومات الكاذبة ونشرها، واستهداف الجماهير والصحفيين بقصد إعادة توجيههم أو تضليلهم، لا سيّما من خلال تقديم «أدلة» على وجود دعم شعبي لشخص أو فكرة أو سياسة ما. رجاءً مراجعة تعريف تكنوبيديا: https://www.techopedia.com/definition/13920/astroturfing

[3] خلال هذه الدراسة، يتم استخدام مصطلح «شركات التواصل عبر الإنترنت» للإشارة إلى الشركات الكبيرة العاملة في مجال محركات البحث ومواقع التواصل الاجتماعي وتطبيقات المراسلة. وهذا ما يسمح بتجنب الإشارة إلى هذه الشركات بتعبير عام على أنها «منصات» والتأكيد على تنوعها. كذلك فإن ذلك مرده إلى أن هذه الشركات ليست مجرد خدمات هيكلية تكنولوجية محايدة أو سالبة، بل مؤسسات ذات مصالح والتزامات وتشكيلات لها تأثير كبير على المعلومات بصفةٍ عامة، والمعلومات المُضلّلة والاتصالات على وجه الخصوص.

المعلومات المُضلّلة والحد منها ومكافحتها من جهة أخرى. على سبيل المثال، كيف تتفاعل مسائل تنظيم المُحتوى وحُريّة التّعبير والتضخيم الخوارزمي للمعلومات المُضلّلة مع بعضها البعض؟

فبموجب قانون حقوق الإنسان، يتمتع التّعبير عن المُحتوى المغلوط بالحماية –شأنه شأن أي تعبير آخر– مع بعض الاستثناءات. فعلى سبيل المثال، وبموجب العهد الدولي الخاص بالحقوق المدنية والسياسية، تخضع أشكال معينة من خطاب الكراهية، والتحريض على العنف، والخطاب الذي يهدد الحياة البشرية (بما في ذلك المعلومات الصحية المغلوطة على نحوٍ خطير) لقيود مشروعة لأسباب تتعلق بحماية حقوق الإنسان الأخرى، أو لأغراض الحفاظ على الصحة العامة. ومع ذلك، وطالما أن التّعبير لا يصل إلى عتبة تقييد المشروع هذه، فإن للأشخاص الحق في التّعبير عن آراء لا أساس لها والقيام بتصريحات لا تعتمد على الوقائع ولا على الأدلة – تتراوح بين الادعاءات بأن «الأرض مسطحة» إلى الآراء القائلة إنّ «الطقس البارد بشكل غير اعتيادي الذي نعيشه اليوم يعني أنّ الاحتباس الحراري المؤدي لظاهرة الاحترار العالمي هو مجرّد أسطورة» –بما في ذلك على وسائل التواصل الاجتماعي (Allan, 2018). من ناحية أخرى، قد يكون من العادل في كثير من الحالات تجريم الأكاذيب المصممة للاحتيال على الناس لسلب أموالهم، أو لتشويه سُمعة شخص ما، أو لمنع الناخبين من الاقتراع، وذلك بموجب القانون الجنائي أو المدني. وكل هذا يجعل من تناول موضوع المعلومات المُضلّلة أكثر تعقيداً من منظور حُريّة التّعبير.

يرتبط التّعبير المعاصر ارتباطاً وثيقاً بالمزيج المكوّن من تكنولوجيات المعلومات وشركات التواصل عبر الإنترنت، والذي، بالاقتران مع تزايد النفاذ إلى إنترنت النطاق الواسع، يسمح بالنشر الفوري للمعلومات على الشبكات العالمية المتاحة لمليارات الأشخاص. وهذا ييسر حُريّة التّعبير ويسمح للمواطنين بفسحة واسعة من وجهات النظر ومصادر المعلومات هي اليوم أكبر بكثير من أي وقت مضى. وفي عالم مقسم بين أغنياء المعلومات وفقراء المعلومات، يُعتبر ذلك نعمةً للأشخاص الذين لم يكونوا مطلعين على المعلومات من قبل. وعلى عكس ذلك، فإن أدوات حُريّة التّعبير هذه أصبحت تُستخدم بشكل متزايد كسلاح من قبل الجهات الفاعلة التي تسعى إلى التلاعب بالرأي العام عن طريق إدخال وتضخيم المحتويات الكاذبة والمُضلّلة في النظام البيئي للمعلومات الموجودة على الإنترنت.

ويمكن أن يؤدي تزايد توافر المعلومات، المقترن بوجود أنظمة إخبارية أكثر تنوعاً، إلى تعزيز تنوّع الأفكار التي يتعرّض لها الأفراد. وفي داخل ذلك البحر الشاسع الذي يُشكّله النظام البيئي للمعلومات المعاصرة، نجد مزودي معلومات موثوق بهم على غرار المنتجين الصحفيين الذين يرقون إلى مستوى الكفاءة المهنية المستقلة، ومراكز البحوث المستقلة، والمنتجين الآخرين للمعلومات المتصلة بالمصلحة العامة الموثوق بها (مثل مقدمي المشورة الصحية من ذوي السمعة الطيبة)، والمعلقين المُطّلعين. ولكن نجد أيضاً كتلةً من اللاعبين الآخرين يتبعون معايير مختلفة لفائدة المصداقية ويعتمدون أخلاقيات ودوافع مختلفة؛ فينتجون تياراً قوياً خطيراً داخل هذا البحر. وبالتالي، قد يشعر المواطنون بالإرهاق بسبب فيض المُحتوى الذي يتعرضون له عبر الإنترنت، فيلجؤون إلى مصادر معلومات زائفة تجذب تحيّزاتهم وتؤكد معتقداتهم أو هُويّاتهم القائمة. ونتيجةً لذلك، بدلاً من أن يكونوا غير مطلعين على المعلومات، يساهمون بطريقة فعّالة في تعريض أنفسهم للمعلومات المُضلّلة أو بطريقة غير مباشرة للمعلومات المغلوطة أو غير الدقيقة.

وقـد أظهـرت الأبحـاث الحديثـة أنّ المعلومـات المُضلّلـة تؤثـر علـى الـدول بدرجـات متفاوتـة (Humprecht, Esser & Van Aelst, 2020) ولا شـك فـي أنّ ارتفـاع الفصـل الأيديولوجـي والاسـتقطاب السياسـي هـو مـن بيـن الدوافـع الرئيسـية وراء ارتفـاع الإنتـاج وانتشـار المعلومـات المُضلّلـة عبـر الإنترنـت فـي بعـض البلـدان (Humprecht, Esser & Van Aelst, 2020).

وعلـى النقيـض مـن ذلـك، تُشـير بحـوث أخـرى إلـى أن اسـتخدام المعلومـات الرقميـة يمكـن أن يـؤدي إلـى التعـرّض لطائفـة أوسـع مـن مصـادر المعلومـات، علـى الرغـم مـن أن ذلـك لا يعنـي بالضـرورة أنّ المُحتـوى بحـد ذاتـه هـو أكثـر تنوعـاً، ولا أنّ المعتقـدات القائمـة هـي بالتالـي متنوعـة. ومـع ذلـك، مـن المعـروف أنّ التعـرض المتكـرر للأكاذيـب يُقلّـل مـن مقاومـة الفـرد للمعلومـات المُضلّلـة. وهـذا مـا ينطبـق أيضـاً علـى التعـرض لمسـتويات عاليـة مـن الرسـائل الشـعبوية (Humprecht, Esser & Van Aelst, 2020).

وعلـى عكـس ذلـك، فـإن القـدرة علـى مقاومـة المعلومـات المُضلّلـة أعلـى فـي البلـدان التـي تعلـو فيهـا الثقـة فـي وسـائل الإعـلام الإخباريـة وتكـون فيهـا رسـائل الخدمـة العامـة عبـر الإعـلام قويـة. كذلـك، قـد يـؤدي انخفـاض ثقـة الجمهـور فـي وسـائل الإعـلام والمؤسسـات الديمقراطيـة إلـى اسـتخدام المعلومـات بشـكل عالـي الانتقائيـة مـن خـلال «غـرف الصـدى» عبـر الإنترنـت التـي تعمـل علـى تضخيـم التّضليـل وتعميـق حالـة الاسـتقطاب.

وبنـاءً علـى ذلـك، هنـاك حاجـة مُلحّـة ليـس فقـط للتصـدي للمعلومـات المُضلّلـة وإنمـا أيضـاً لاتخـاذ خُطـوات نحـو إعـادة بنـاء العقـد الاجتماعـي وثقـة الجمهـور بالقواعـد والمعاييـر الدوليـة المتفـق عليهـا؛ أي تعزيـز المؤسسـات الديمقراطيـة وتعزيـز التماسـك الاجتماعـي بصفـةٍ خاصـة فـي المجتمعـات شـديدة الانقسـام، واعتمـاد أسـاليب بنـاء جُسـور الحـوار للتوجُّـه للجماعـات والجهـات المتحصنـة بالإنترنـت.

لـذا، لابـد مـن دراسـة الاسـتجابات المتنوعـة المتصديـة للمعلومـات المُضلّلـة علـى الصعيـد العالمـي، ووضـع أُطُـر تسـاعد علـى فهـم وتقييـم هـذه الاسـتجابات مـن منظـور حُريّـة التّعبيـر. هـذا هـو الهـدف الأساسـي التـي تسـعى إلـى تحقيقـه هـذه الدراسـة، والتـي أُجريـت البحـوث لهـا بيـن أيلول/سـبتمبر 2019 وتمـوز/يوليـو 2020.

قبـل أن نبـدأ بفـك رمـوز هـذه الأبعـاد وتشـريحها، مـن الضـروري تحديـد المعاييـر المُتبعـة بطـول الدراسـة، وشـرح المصطلحـات الرئيسـية المسـتخدمة، والنظـر فـي بعـض الأمثلـة المتعلقـة بالتضليـل عبـر الإنترنـت، وعلاقتـه بالدعايـة السـلبية أو «البروباغنـدا»، والمعلومـات المغلوطـة وخطـاب الكراهيـة.

1.1 تقنيات التّضليل عبر الإنترنت

إنّ انتشار المعلومات المُضلّلة عبر الإنترنت يطرح أسئلةً خطيرةً حول الدور الذي يلعبه كلّ من البحث على الإنترنت ووسائل التواصل الاجتماعي والرسائل الاجتماعية والإنترنت عموماً في الديمقراطيات المعاصرة. وتكثر الأمثلة على المعلومات الرقمية المُضلّلة، بدءً من التدخل في الانتخابات، ووصولاً إلى المعلومات المُضلّلة الطبية (مثلاً: حول التلقيح[4] وجائحة فيروس كورونا[5]) والتي غالباً ما تنطوي على تهديدات بالأذى الجسدي، وخطر على الخصوصية، والإضرار بسمعة الأفراد والصحة العامة.

وفي حين أنّه غالباً ما تتم دراسة المعلومات المُضلّلة في ما يخص «تويتر»، و«فيس بوك»، و«يوتيوب»، فإن هذه المعلومات موجودة أيضاً على العديد من المنصات الاجتماعية الأخرى (على سبيل المثال: رديت، وانستغرام، وتيك توك، وفور تشان chan4، وبينتيرست[6])، كما وعلى تطبيقات المراسلة (على سبيل المثال: واتساب، وتلغرام، وسناب تشات، وآي–مسدج iMessage)، وعلى محركات البحث على الإنترنت (على سبيل المثال: غوغل). وهناك أيضاً مواقع مخصصة للمعلومات المُضلّلة (مثل: Infowars وQ-anon). بالإضافة إلى ذلك، ستتم عند الاقتضاء الإشارة إلى الجهات الفاعلة والوسائط الأُخرى (كمزودي خدمات الإنترنت ومقدمي خدمات الحوسبة السحابية). وعلى الرغم من أن الدراسة كانت شاملةً عند كتابة هذا التقرير الجامع، فإنها تقرّ أيضاً بالحاجة إلى مواصلة البحث في الآليات الناشئة لنشر المعلومات المُضلّلة، وفي المنصات الاجتماعية الجديدة التي تتطوّر بشكل هائل، بما في ذلك تلك التي يُنظر إليها بشكل رئيسي على أنها مواقع ترفيهية واجتماعية (على غِرار: تيك توك) وليس كمنصات لأغراض سياسية وما إلى هنالك.

وفي حين أن أبرز منتجي وقنوات المعلومات المُضلّلة هم غالباً اللاعبون السياسيُّون والدول (Brennan et al 2020, Billings 2020, Bradshaw & Howard 2019) فإنّ هذا التقرير لا يُركّز على مصادر المعلومات المُضلّلة ولاعبيها، ولكن بالأحرى على الاستجابات للمعلومات المُضلّلة في جميع أنحاء العالم. ومن بين هذه الاستجابات، تضطلع الدول والجهات الفاعلة السياسية بدور حاسم في وقف المعلومات المُضلّلة من المنبع – بما يتضمن من داخل «مصادرها». ويتم تقييم أهميتها بصفة خاصة في مجال الاستجابات المتعلقة بالرسائل المضادة، والتشريعات والسياسات، والانتخابات والتدخلات المعيارية.

وتُناضل عدة جهات عادية منتجة للأخبار –على الإنترنت وخارجه– من أجل الحفاظ على مركزها كمرجع لمن يبحث عن معلومات جديرة بالثقة ضمن بيئة الاتصالات الأوسع هذه. وقد أصبحت الوسائل الإخبارية أيضاً من القنوات الناقلة للمعلومات المُضلّلة في بعض الحالات، وذلك بسبب ضعف معايير التحقُّق، وتلاعب جهات خارجية بها، وحتى تواطؤها في بعض الأحيان (كوسائل الإعلام شديدة التحزُّب).

ومع ذلك، فإنّ الدور العام الذي تضطلع به وسائل الإعلام المستقلة الناقدة في إضفاء الشرعية على الأخبار ووضع الأجندات المجتمعية يجعلها أيضاً هدفاً رئيسياً لموردي المعلومات المُضلّلة. وفي حالة حملات التّضليل الإعلامية المنظمة، غالباً ما تستهدف الهجمات مصادر المعلومات المشروعة والموثوقة

4 https://firstdraftnews.org/long-form-article/first-draft-case-study-understanding-the-impact-ofpolio-vaccine-disinformation-in-pakistan/

5 https://www.poynter.org/fact-checking/2020/coronavirus-fact-checkers-from-30-countries-arefighting-3-waves-of-misinformation/%20

6 https://medium.com/dfrlab/trudeaus-and-trudeaunts-memes-have-an-impact-during-canadianelections-4c842574dedc

-على غرار وسائل الإعلام الإخبارية والصحفيين الموثوق بهم وذلك من خلال أعمال القرصنة والتخريب وغيرها من أساليب الترهيب والمراقبة بغية تطبيق استراتيجية شاملة لنشر التّضليل وتحقيق أهداف أوسع نطاقاً. ولا شك في أنّ العديد من حملات التّضليل المنظمة هي حملات تشنها الحكومات و/أو جهات فاعلة سياسية وجيوسياسية؛ مما يدفعنا إلى فهم دور الدولة في الاستجابة للمعلومات المُضلّلة. ومع ذلك، فإن الغرض الرئيسي من هذا التقرير البحثي هو تفصيل الطرق المتنوعة للاستجابة لأزمة المعلومات المُضلّلة العالمية، لا تقييم المُحرّضين والوكلاء ودوافعهم.

لا يعرف التّضليل حدوداً ضمن بيئة المعلومات، بل يتسرّب إلى قنوات تواصل متعددة بفعل تصميمها، أو بفعل إعادة النشر والتضخيم الذي تغذيه هيكليات شبكات النظير للنظير والصديق للصديق.

وفي ما يخص المُحتوى، تم تحديد ثلاثة أنواع رئيسية للمعلومات المُضلّلة في هذه الدراسة، استناداً إلى طريقة عرض المُحتوى (أي نص مكتوب، أو صورة، أو مقطع فيديو، أو صوت، أو مُحتوى مُختلط) والطريقة التي تم بناؤه أو التلاعب به:

1. **البنى السردية العاطفية و«الميمات»**: وهي ادعاءات كاذبة وسرديات مكتوبة[7] غالباً (ولكن ليس دائماً) ما تقوم بمزج كلمات عاطفية قوية وأكاذيب و/أو معلومات غير مكتملة وآراء شخصية ببعض العناصر الحقيقية. يصعب الكشف عن هذه البنى على تطبيقات الرسائل المغلقة ويتم تطبيقها على مجموعة كبيرة من المُحتوى بدءً من «الأخبار الزائفة» إلى الإعلانات السياسية الإشكالية.

   - **السرديات المغلوطة و/أو المُضلّلة** التي تُحاكي أشكالاً معينةً على غِرار كتابة الأخبار أو البرامج الوثائقية، والتي عادةً ما تمزج الادّعاءات المكتوبة الخاطئة أو المعلومات غير المكتملة بالآراء الشخصية، والصور و/أو مقاطع الفيديو و/أو المقاطع الصوتية، والتي يمكن أن تكون بحد ذاتها خاطئة أو تم التلاعب بها، أو أُخِذَت في غير سياقها. ويتم استخدام محتويات من مواقع أخرى بشكل مخصص أحياناً لخلق انطباع عام مضلل أنه مجمّع أخبار محايد.

   - **السرديات العاطفية** التي تترافق بآراء شخصية وصور و/أو مقاطع فيديو ومقاطع صوتية قد تكون غير حقيقية، أو تم التلاعب بها، أو مأخوذة في غير سياقها، والتي تسعى أيضاً إلى إملاء تفسيرات لمعلومات معينة، على سبيل المثال: للتقليل من أهميتها، أو لتشويه سمعة المصدر.

2. **الصور ومقاطع الفيديو[8] ومقاطع الصوت المُركَّبة[9]** التي تم تعديلها أو اختلاقها أو تجريدها من سياقها والتي تُستخدم لخلق جوّ من الارتباك وانعدام الثقة المعمم و/أو إثارة عواطف قوية من

[7] قاعدة بيانات تضم أكثر من 6000 ادعاء ورواية كاذبة حول جائحة فيروس كورونا؛ خضعت لعملية التدقيق في الحقائق من أكثر من 60 دولة: https://www.poynter.org/coronavirusfactsalliance/

[8] إنّ الصور أو مقاطع الفيديو المأخوذة في غير إطارها هي محتويات حقيقية موجودة مسبقاً، تم إعادة استخدامها كجزء من سردية كاذبة لنشر المعلومات المُضلّلة. على سبيل المثال، تم استخدام شريط فيديو قديم لأشخاص يُصلّون في تغريدة لليمين المتطرف تدّعي أنّ المسلمين ينتهكون قواعد التباعد الاجتماعي في ظل جائحة كورونا؛ والحدث نفسه صحيح وقديم أما توظيفه في ظل الجائحة فغير صحيح.

[9] رجاءً مراجعة التعريف أدناه.

خلال «الميمات» واسعة الانتشار أو الروايات الكاذبة. وتنطبق هذه أيضاً على مجموعة واسعة من المحتويات بدءً من الدعاية السياسية إلى الإعلانات الكاذبة. ومن بين هذه التقنيات نميّز بين الفئات التالية:

- **الصور ومقاطع الفيديو** المأخوذة في غير سياقها من دون تعديل أو تقريباً من دون تعديل مع مستوىً عالٍ من التشابه، والتي غالباً ما تشمل النُّسخ التي تستخدم لأغراض خلق عناوين تخطف الانتباه clickbait؛

- **المقاطع الصوتية والصور ومقاطع الفيديو** التي تم تقسيمها إلى جزء أو عدة أجزاء من المقطع الصوتي أو مقطع الفيديو الأصلي، أو التي تم تعديلها لإزالة الطابع الزمني timestamp من لقطات كاميرات المراقبة على سبيل المثال. وتسمى أيضاً بالمقاطع «ذات التزييف البسيط»؛

- **مقاطع الفيديو** المنتجة عن قصد، وعلى سبيل المثال تلك التي تنتجها شركة إنتاج الفيديو؛

- **الصور المُنشأة عبر الحاسوب** Computer-Generated Imagery (CGI) بما في ذلك الصور ذات التزييف الفج deepfakes (صور/مقاطع فيديو كاذبة يتم استحداثها بواسطة الذكاء الاصطناعي) والتي يتم إنتاجها بالكامل بواسطة الحاسوب، أو التي تمزج صور/مقاطع مصورة/مقاطع صوتية موجودة مسبقاً.

- **المقاطع الصوتية المركبة**: إنّ تركيب الكلام، الذي يستخدم البرمجيات المتقدمة لخلق نموذج لصوت شخص ما هو نوع جديد نسبياً من التزييف العميق. وهذا ينطوي على نسخ صوت يمكنه أن ينطق بنص باستخدام الإيقاع والنبرة كما الشخصية المُنتَحَلة. وتسمح بعض التقنيات (على غِرار: Modulate.ai) للمستخدمين بخلق أصوات مركبة بالكامل قادرة على تقليد أيّ جنس أو عُمر. (Innovation and Ethics Data for Centre 2019)

3. **المواقع الزائفة ومجموعات البيانات المُغرِضة**، بما في ذلك المصادر الزائفة ومجموعات البيانات التي تم التلاعب بها والمواقع الحكومية أو مواقع الشركات الزائفة (Trend Micro, 2020). تتضمن هذه الفئة أيضاً مواقع على الإنترنت تستخدم أسماءً تجعلها تبدو وكأنها وسائل إعلام إخبارية؛ وتنشر معلومات قابلة للتصديق على شكل قصص إخبارية، على سبيل المثال: الإبلاغ عن حالات زائفة للعدوى بفيروس كورونا (Thompson, 2020).

وتُسخّر هذه الطرق المختلفة للتضليل في إطار مجموعة من الممارسات الضارة، بما في ذلك على سبيل المثال لا الحصر:

- الحملات الإعلامية التي ترعاها الدولة؛
- الدعاية الحكومية (أو المناهضة للحكومة)/الدعاية السياسية الأخرى؛
- قيام القادة السياسيون بإنتاج مُحتوى كاذب ومضلل وتضخيمه؛

- العناوين المثيرة للانتباه[10] clickbait؛
- إعلانات كاذبة أو مضللة؛ تكون على سبيل المثال متصلة بالسياسة وإعلانات الوظائف؛
- انتحال شخصية وسائل الإعلام الموثوقة، أو منظمات التحقُّق من الوقائع، أو أفراد أو حكومات (مواقع و/أو حسابات زائفة على وسائل التواصل الاجتماعي، روبوتات الإنترنت)؛
- حملات الدعاية الشعبية الزائفة Astroturfing؛
- المنتجات والمراجعات الزائفة؛
- المعلومات المغلوطة المضادة للّقاحات، وحول فيروس كورونا، وغيرها من المعلومات المتعلقة بالصحة والطب والرفاه؛
- زرع الشك[11] Gaslighting؛
- الهويات والسلوكيات المُفتعلة.

يُستثنى من قائمة ممارسات التواصل هذه الهجاء السافر والسخرية، على الرغم من أن هاتين الممارستين قد تؤديان في بعض الحالات إلى التّضليل وإلحاق الضرر بالمواطنين الذين يفتقرون إلى الدراية الإعلامية والمعلوماتية التي تسمح لهم بتمييز السُّخرية عن الواقع. ويمكن للهجاء والسُّخرية، في الواقع، أن يكونا أداةً فعالةً للتصدي للمعلومات المُضلّلة عن طريق تسليط الضوء على العناصر غير المنطقية فيها (مَن ينتجها وينشرها) بطرق فعّالة وجذابة. ومع ذلك، ينبغي ألا تُعتبر ممارسات التواصل هذه من المعلومات المُضلّلة.

[10] تدوينة مصممة لإثارة رد عاطفي في صفوف من يقرأها (مثل: الغضب والتعاطف والحزن والخوف)، وبالتالي تدفع بالمستخدم إلى مزيد من المشاركة (مثلاً عبر «النقر») عبر الضغط على الرابط لفتح صفحة الإنترنت؛ والتي بدورها تخلق مُشاهدات للإعلانات وإيرادات لمالك الموقع. السمة المميزة للعناوين الخاطفة أو clickbait هي أن مُحتوى الخبر لا يعكس العنوان، وهذا يعني أن «الناقر» قد وقع في الفخ الذي نُصب له عبر الطُّعم من دون أن يفي المقال بتوقعاته.

[11] شكل من أشكال التلاعب النفسي: https://dictionary.cambridge.org/dictionary/english/gaslighting

2.1 التّعريفات والنّطاق

هناك العديد من التعريفات المختلفة والمتباينة إلى حدٍّ ما للمعلومات المُضلّلة وما إذا كانت ترتبط بالمعلومات المغلوطة أو لا، وبأي طريقة. وعادةً ما يشترك المفهومان في سمة الزيف كمعيار أساسي، مما يؤدي إلى استخدام مصطلحي المعلومات المغلوطة والمعلومات المُضلّلة بشكل مترادف ومتبادل (على سبيل المثال في Alaphilippe, Bontcheva et al., 2018b).

أما قاموس أكسفورد الإنكليزي [12] (OED) فيميّز بين المصطلحين على أساس أنّ الأول مرتبط بنية الخداع والآخر بنية التّضليل (على الرغم من أنه ليس من الواضح كيف تختلف هاتان النيتان):

- **معلومات مغلوطة**: معلومات خاطئة أو غير دقيقة، خاصةً تلك التي يُقصد منها الخداع عمداً.
- **معلومات مضللة**: معلومات خاطئة يُقصد منها التّضليل، وخاصةً الدعاية التي تصدرها منظمة حكومية ضد سلطة أو جهة إعلامية منافسة لها.

ويربط هذا التعريف أيضاً أحد المصطلحين (المعلومات المُضلّلة) بجهة فاعلة معينة (حكومية)، مما قد يضيّق نطاق المسؤولية. كذلك عرّف آخرون المعلومات المُضلّلة تحديداً في سياق الانتخابات، على أنها «مُحتوى تم اختلاقه عمداً بهدف تخريب العمليات الانتخابية» (Giglietto et al., 2016). ولا شك في أنّ هذا التعريف هو أيضاً ضيق للغاية بالنسبة إلى المجموعة الواسعة من المعلومات المُضلّلة التي تم تناولها في هذه الدراسة.

وثمة منظور آخر يرد في توصيات التقرير الذي أعده فريق الخبراء رفيع المستوى التابع للاتحاد الأوروبي المعني بالأخبار الزائفة والمعلومات المُضلّلة على الإنترنت، والذي يتضمن إشارات إلى المعلومات المُضلّلة التي تسعى للربح كجزء مما يشمله المصطلح:

> **”إنّ المعلومات المُضلّلة... تشمل جميع أشكال المعلومات الكاذبة أو غير الدقيقة أو المُضلّلة المُصمَّمة والمُقدَّمة والمُروَّج لها عن قصد بهدف التسبُّب بالضرر العام أو [تحقيق] الربح (Buning et al., 2018)“**

ولكن اشتراط وجود نية في الربح يحُد من نطاق المعلومات المُضلّلة. فعلى سبيل المثال، هدف الهجاء هو تحقيق الربح لمنتجي التلفزيون والمجلات الساخرة، إلا أنه من الصعب اعتبار الهجاء في حد ذاته من المعلومات المُضلّلة.

إنّ الإطار النظري لاضطراب المعلومات المعتمد على نطاق واسع (Wardle, 2017a; Wardle & Derakhshan, 2017) يُميّز بين المعلومات المغلوطة والمعلومات المُضلّلة على النحو التالي:

[12] https://www.lexico.com/definition/misinformation

- **المعلومات المغلوطة:** معلومات خاطئة يتم مشاركتها عن غير قصد، من دون النية في إلحاق الضرر.
- **المعلومات المُضلّلة:** عملية تهدف إلى التسبُّب في ضرر عن طريق مشاركة معلومات خاطئة عمداً.

ويمكن تمثيل المعايير الواردة في هذا الإطار على النحو التالي:

	إدراك زيف المعلومات	القصد الكامن
معلومات مُضللة	مُدرك	مُسيء
معلومات مغلوطة	غير مُدرك (غير مقصود)	جيّد/محايد

تتماشى هذه التعاريف بشكل عام مع تعاريف قاموس كامبريدج الإنكليزي[13]، الذي يعرّف المعلومات المُضلّلة على أنها تتضمن نيةً للخداع، في حين أنّ المعلومات المغلوطة أكثر غموضاً.

وتشترك معظم التعاريف في سمة القصد أو النية في ما يتعلق بالضرر (في قاموس أوكسفورد يشار ضمنياً إلى أن ممارسات الخداع ومحاولات التّضليل هي جميعها ممارسات سلبية).

وفي الوقت نفسه، فإن التمييز بين النوعين بناءً على النية (وإدراك الزيف) أمر معقد لأنّه قد يكون من الصعب على الخوارزميات، كما وعلى المتلقين من البشر، تحديد الدافع ومصدر المعلومات أو الجهة المضخمة في كثير من الأحيان (Jack, 2017; Zubiaga et al., 2016). وهناك أيضاً خطر اعتماد «الافتراض المانوي» حول هُويّة «الجهة الفاعلة السيئة»، وهو افتراض قد يؤدي إلى تبسيط الموضوع بشكل كبير، وينطوي على تفسيرات إشكاليةً تتسم بالذاتية أو التحزب للمصالح المقصود الإضرار بها ولأي شخص أو جهة هي تابعة.

وما يظهر نتيجةً لذلك هو التحدي المتمثل في وضع القصد وإدراك الزيف في صميم تعريف ما ينبغي اعتباره معلومات مضللة، في مواجهة ظاهرة أوسع نطاقاً من المُحتوى المغلوط أو المضلل. هذا يفسر جزئياً لماذا يقارب بعض الكُتّاب (على سبيل المثال، Francois, 2019) القضية لا من باب النية (أو إدراك الوكيل) في المقام الأول ولكن من باب التركيز على السلوكيات المرئية كالعمليات المُعد لها التي تتضمن استخدام روبوتات الإنترنت (والتي قد تشير إلى جود نية ضارة وإدراك للزيف). وقد تشير السلوكيات المنظمة (بما في ذلك من قبل جهات فاعلة لا أساس لها) إلى وجود مُحتوى مغلوط ولكنّ الأكاذيب الضارة قد تنتشر أيضاً من دون تضخيم خاص، وكلها غالباً ما تنشأ من جهات فاعلة أصيلة على غرار المشاهير والسياسيين، كما هو مبين في البحوث (e.g. Brennen et al 2020; Satariano & Tsang 2019). في الوقت نفسه، يمكن تعميم المُحتوى الصادق من خلال سلوكيات وجهات فاعلة مختلفة كجزء من حملة لنشر المعلومات أو لمُكافحة المعلومات المُضلّلة، وهذا ما يختلف عما يقصد به في مصطلح «المعلومات الضارة» الذي يعتمده (Derakhshan & Wardle, 2017) في ما يخص المُحتوى المأخوذ في غير سياقه أو الموضوع في إطار مزيّف. لذلك، سيكون مقيّداً لهذه الدراسة تضييق نطاق معالجتها للمعلومات المُضلّلة وكأنها تقتصر على السلوكيات (حيث أنها غالباً إشارة إلى وجود تحديات).

13 https://dictionary.cambridge.org/dictionary/english/misinformation

ولأنّ هـذه الدراسـة تسـعى، مـن جانبهـا، إلـى تغطيـة طائفـة واسـعة مـن الاسـتجابات مـن جميـع أنحـاء العالـم، فإنهـا تتجنـب اتبـاع مقاربـة ضيقـة فـي تعريـف المعلومـات المُضلّلـة. لذلـك، فإنهـا تسـتخدم مصطلـح المعلومـات المُضلّلـة بشـكل عـام لوصـف المُحتـوى المغلـوط أو المُضلّـل الـذي قـد يسـبب ضـرراً محـدداً -بغض النظـر عـن الدوافـع أو الإدراك أو السُّـلوكيات. وقـد يكـون هـذا الضـرر، علـى سـبيل المثـال، إضـراراً بالديمقراطيـة والصحـة والأقليـات والمجتمعـات المحرومـة، والتحديـات المناخيـة، وحُريّـة التّعبيـر. وهنـا أيضـاً تعتمـد المقاربـة العمليـة لمـا يُشـكّل معلومـات مضللـة (وبالتالـي الاسـتجابات لهـذه الظاهـرة) علـى وجـود الزيـف والأثـر السـلبي المحتمـل علـى الجهـات المسـتهدفة، لا علـى قصـد أو إدراك أو سـلوك الجهـات المنتجـة أو الموزعـة لهـذه المعلومـات. إضافـةً إلـى ذلـك، إذا اعتبرنـا أنّ المعلومـات المغلوطـة بمعناهـا الضيـق هـي مشـاركة غيـر مقصـودة للمعلومـات الخاطئـة مـن دون قصـد التسـبب فـي ضـرر، يبقـى أنّ مُحتـوى هـذه المعلومـات يديـن بأصلـه إلـى أعمـال مقصـودة مـن قبـل آخريـن تهـدف إلـى تضليـل المواطنيـن بقصـد إلحـاق الضـرر. وإنّ الاعتـراف «بمصـدر» هـذا الضـرر هـو سـبب قـوي دفعنـا إلـى اعتمـاد مصطلـح «المعلومـات المُضلّلـة» كمصطلـح عـام فـي هـذه الدراسـة بـدلاً مـن مصطلـح «المعلومـات المغلوطـة».

ولا ينبغـي أن تكـون هـذه المقاربـة اختزاليـةً بطبيعتهـا؛ بمعنـى أن تصنّـف كافة المعلومـات علـى أنها إمّـا معلومـات مضللـة محتملـة، وإمّـا معلومـات (صحيحـة) لا تلحـق الأذى. فـالآراء التـي تعكـس القيـم والمواقـف هـي من الأمثلة علـى المُحتـوى الـذي لا يمكـن تصنيفـه علـى أنـه صحيـح أو خاطـئ. كذلـك فالعلـم والسياسـة همـا مسـألتان متغيرتـان وتتطـوران بمـرور الزمـن وقـد لا تتوافقـان، فـي البدايـة علـى الأقـل، مـع هـذا التقييـم الثنائـي. ومـن جانبهـا، فـإن المعلومـات المُضلّلـة، بطبيعتهـا، تُضفـي صفـة «الصحيـح» ليـس فقـط علـى الأكاذيـب ولكـن أيضـاً وفـي كثيـر مـن الأحيـان علـى مـا هـو مجهـول أو غيـر معلـوم، فيمـا تسـعى غالبـاً إلـى تشـويه سـمعة المُحتـوى الـذي أثبتـت صحتـه عبـر اعتبـاره «كاذبـاً» -علـى غـرار توافـق الآراء العلمـي المهـول بشـأن مسـألة تغيُّـر المنـاخ. وإن وجـود مناطـق رماديـة فعليّـة يفسـر وجـود مخاطـر فـي أي خُطـوات تُتخـذ لمواجهـة المعلومـات المُضلّلـة إذا مـا تجاهلـت الحيـز الواسـع الموجـود بيـن الحقائـق المثبتـة والأكاذيـب التـي تم تبيـان زيفهـا. ويمكـن لهـذه التدابيـر أن تخنـق النقـاش المشـروع وأشـكال التّعبيـر الأخـرى التـي نحتـاج إليهـا لتقييـم صحـة مُحتـوى معيـن علـى مر الزمـن.

وإنّ اسـتخدام المعلومـات المُضلّلـة كمصطلـح عـام لتقييـم الاسـتجابات التـي تسـعى إلـى التصـدي للمُحتـوى المغلـوط لا يحـول دون الاعتـراف بـأن هـذه الاسـتجابات قـد تختلـف باختـلاف دوافـع (إن كانـت ماليـةً أو سياسـيةً أو إيديولوجيـةً أو متعلقـةً بالأحـوال الشـخصية ومـا إلـى ذلـك) أو سـلوكيات الجهـات المُحرّضـة أو الفاعلـة فـي مجـال التّضليـل. والتثقيـف مثـلاً هـو عـلاج جزئـي للمعلومـات المغلوطـة (عندمـا تشـير إلـى إنتـاج أو تـداول غيـر مقصـود للأكاذيـب مـن دون سـوء نيـة أو إدراك لزيـف المُحتـوى) فـي حيـن أنّ سـنّ الأنظمـة لمنـع كسـب المـال عـن طريـق الرسـائل المخادعـة هـو مـن الطـرق التـي تحـد مـن توفيـر المعلومـات المُضلّلـة (وهنـا اسـتخدام للمصطلـح بمعنـاه الضيـق للإشـارة إلـى الزيـف الواعـي والمتعمـد). وقـد يكـون تشـويه الحقائـق والخـداع مـن الممارسـات الأكثـر انتشـاراً فـي السـياقات السياسـية والانتخابيـة، فـي حيـن أنّ المعلومـات المغلوطـة (بالمعنـى الضيـق) تنتشـر أكثـر فـي حالـة الحمـلات المناهضـة للتلقيـح. وبالتالـي، فـإن نظريـة التغييـر الكامنـة التـي تنطـوي عليهـا اسـتجابة معينـة، غالبـاً مـا ترتبـط بافتراضـات بشـأن النيـة والسـلوكيات ذات الصلـة. ومـع ذلـك، وخاصـةً فـي سـياق الانتخابـات والاسـتفتاءات والجائحـات علـى غـرار جائحـة فيـروس كورونـا، فـإن التأثيـر الضـار للمُحتـوى المغلـوط، بغـض النظـر عـن النوايـا، وبغـض النظـر عـن مجموعـة السـلوكيات الكامنـة وراءه، قـد يكـون هـو نفسـه حيـث أنـه يـؤدي إلـى إضعـاف الأفـراد وإلـى آثـار خطيـرة. لذلـك، يجـب تصميـم التدخـلات بشـكلٍ مناسـب.

ونظـراً إلـى اختصـاص هـذه الدراسـة، فمـن المنطقـي أن يسـتخدم الإطـار الدلالـي مصطلـح «المعلومـات المُضلّلـة» كعبـارة عامـة تُغطّـي المعلومـات المغلوطـة فـي مُحتـوى معيّـن (التـي تشـمل الرسـائل المُضلّلـة) التـي ترتبـط بضـرر مجتمعـي محتمـل (مثـل: الآثـار السـلبية علـى حقـوق الإنسـان والصحـة العامـة والتنميـة المسـتدامة). وهـذا مـا يسـمح باسـتعراض المجموعـة الواسـعة مـن الاسـتجابات للمعلومـات المُضلّلـة المعتمدة فـي جميـع أنحـاء العالـم. وبالتالـي، لا يتمثـل القصـد فـي وضـع تعريـف آخـر لماهيـة المعلومـات المُضلّلـة، بـل فـي توفيـر مفهـوم شـامل واسـع للمسـألة قيـد البحـث والتحليـل. وعلـى هـذا الأسـاس الواسـع، يحـرص البحـث علـى الإشـارة، حسـب الاقتضـاء، إلـى الطريقـة التـي تفسـر بهـا هـذه الظاهـرة مـن قبـل مختلـف أصحـاب المصلحـة الذيـن يسـتجيبون للمعلومـات المُضلّلـة، ضمنـاً أو صراحـةً -ضمـن الاسـتجابة المعينـة قيـد المناقشـة.

ومـن خـلال اعتمـاد هـذه المقاربـة، فـإنّ هـذه الدراسـة قـادرة علـى إظهـار كيـف تتم مواجهـة ظاهـرة المعلومـات المُضلّلـة، وهـو أمـر مُعقّـد، عبـر اسـتجابات متباينـة فـي جميـع أنحـاء العالـم، وأثـر هـذه الاسـتجابات علـى حُريّـة التّعبيـر والتنميـة المسـتدامة. وفـي الوقـت نفسـه؛ لابـدّ مـن تسـليط الضـوء علـى كيفيـة تعريـف مـا لا ينـدرج تحـت المعلومـات المضللـة.

وبنـاءً علـى ذلـك، لا ينبغـي اختـزال المعلومـات المُضلّلـة بالزيـف الـذي لـه ضـرر محتمـل فقـط فـي إطـار الأخبـار (كمـا يُعنـى بـه ضمنـاً فـي تسـمية «الأخبـار الزائفـة»). كذلـك وكمـا هـو موضـح أدنـاه، ينبغـي أيضـاً عـدم الخلـط بيـن المعلومـات المُضلّلـة وبيـن الدعايـة أو خطـاب الكراهيـة.

3.1 وضع تصوُّر لدورة حياة المعلومات المُضلّلة

لفهم الاستجابات التي تسعى إلى التصدي للتضليل على الإنترنت بفعالية، من الضروري التركيز ليس فقط على الرسالة نفسها وصحتها، ولكن أيضاً على التحقيق في جميع جوانب دورة حياة المعلومات المُضلّلة، بما في ذلك انتشارها وتأثيراتها على المُتلقين المُستهدفين.

ويُطلق على أحد الأُطُر المفاهيمية إطار «ABC»، الذي يميّز بين الجهات الفاعلة والسلوك والمُحتوى. ويسعى الإطار إلى التركيز على الجهات الفاعلة «المتلاعبة» التي تشارك في التّضليل وهي على علم بالموضوع. كما ويركّز أيضاً على السُّلوك المُخادع المُفتعل على الشبكات كما في العمليات الإعلامية وعلى المُحتوى الذي ينشر الأكاذيب (باستخدام تنسيقات إعلامية تم التلاعب بها) أو أموراً قد تكون واقعيةً ولكنها تسعى إلى التحريض (Francois, 2019; Annenberg Public Policy Center, 2020). والدافع هنا هو تعزيز الاستجابات كي لا تتم مُحاربة محتويات قد تكون «بغيضةً» ولكنها تبقى تعبيراً لابدّ من أن يحظى بالحماية في المجتمعات الديمقراطية. ولذلك، فإنّ الإطار يطرح مسألة ما إذا كان من الأفضل أن تركز الاستجابات على ألف A وباء B أكثر من على جيم C.

أما إطار "AMI" فهو مقاربة مفاهيمية أخرى (Wardle & Derakhshan, 2017) يُميّز بين:

- **الوكلاء**، أي مؤلّفي أو موزّعي المعلومات الخاطئة ودوافعهم؛
- **الرسالة**، أي المُحتوى غير الصحيح و/أو الذي تم التلاعب به والذي يجري نشره، وطريقة التّعبير عنه، والتقنيات المستخدمة لتعزيز مصداقيته؛
- **الجهات المفسرة (أو الجهات المستهدفة)**، أي مَن تستهدفهم حملة المعلومات المُضلّلة وتؤثّر على معتقداتهم وتصرفاتهم.

في هذه الدراسة، تم تكييف هذين الإطارين وجمعهما لتشكيل إطار جديد يعكس عنصرين آخرين يلقيان بالمزيد من الضوء على الوكلاء والسلوكيات والجهات الناقلة للمعلومات المُضلّلة:

- المُحرّضون الأصليون على التّضليل، الذين قد يكونون مختلفين عن الوكلاء. هم الجهات الفاعلة التي تقف وراء عملية إنتاج وتوزيع المُحتوى، والذين غالباً ما يقومون بتسخير هذه العملية ودفع تكاليفها. ولذلك فإنّ هذه الجهات هي المصدر الحقيقي والمستفيد الأساسي من هذه المعلومات المُضلّلة. وفي بعض الحالات، قد يكون المُحرّضون هم أنفسهم الوكلاء المنفذين الفعليين، ولكن في الكثير من الحالات واسعة النطاق قد يكون الوكلاء مؤيّدين أو مُتعاقدين طوعيين أو مأجورين أو مُشاركين عن غير قصد. ومع ذلك، فإنّ وظائف المُحرّضين والوُكلاء مُختلفة.
- الوسائط التي هي مِعْوَل أساسي في نقل الرسالة (مثل: مواقع وتطبيقات التواصل الاجتماعي) –مما يدفع بنا إلى الاهتمام بالدور الرئيسي الذي تؤدّيه في نشر ومُكافحة المعلومات المُضلّلة، وكيف يمكن لأنظمة هذه الوسائط أن تمكّن –أو تُعطّل– المُحتوى والجهات الفاعلة والسلوكيات الضالعة في عملية التّضليل.

يمكن وصف هذا الإطار الجامع بالإشارة إلى التالي **1. المُحرّضون 2. الوكلاء 3. الرسائل 4. الوسائط 5. الجهات المستهدفة/المفسرة** – ولذلك يعرف الإطار بمختصر IAMIT. تهدف هذه المقاربة إلى تصوير دورة الحياة الكاملة للمعلومات المُضلّلة – من التحريض والإنتاج إلى وسائل النشر إلى التأثير على أرض الواقع، وذلك من خلال الإجابة عن الأسئلة التالية:

1. المُحرّضون:

- من هم المُحرّضون والمُستفيدون المباشرون وغير المباشرين؟ ما هي علاقتهم بالوكيل (الوكلاء) (المذكور أو المذكورين أدناه)؟ لماذا يتم نشر المعلومات المُضلّلة؟ ما هو الدافع السياسي أو المالي مثلاً؟ هل أن الهدف هو تعزيز موقع معين أو غيرية في غير مكانها؟ ما هي الدوافع الأيديولوجية.. إلخ؟ ويشمل ذلك النظر، حيثما كان ذلك ممكناً، في ما إذا كانت هناك نية في الإضرار و/أو التّضليل.

2. الوكلاء:

- من الذي يقوم فعلياً من الناحية التشغيلية بإنتاج ونشر المعلومات المُضلّلة؟ ويثير هذا السؤال قضايا تتعلّق بإسناد الفاعل (في ما يختص بالهوية الأصيلة)، ونوع الفاعل («مؤثر»، فرد، مسؤول، مجموعة، شركة، مؤسسة)، ومستوى التنظيم والموارد المستعملة، ومستوى التشغيل الآلي «الأُوتوماتيكي». وهنا يظهر استخدام لبعض السلوكيات، مثل: استخدام بعض التقنيات على غرار روبوتات الإنترنت bots والشبكات الدُّمَى sock puppet networks والهُويّات الزائفة.

3. الرسائل:

- ما الذي يجري نشره؟ من الأمثلة على ذلك الادعاءات أو السرديات الكاذبة، والصور ومقاطع الفيديو المأخوذة خارج سياقها أو التي تم تعديلها لغايات الغش والتّدليس، والصُّور الزائفة بفجاجة deep fakes، وما إلى ذلك. هل تشمل الاستجابات فئات تنطوي على معلومات مُضلّلة (على سبيل المثال: المُحتوى السياسي/الانتخابي)؟ ما الذي يُشكّل رسائل قد تكون ضارةً أو هي ضارة بالفعل أو هي ذات ضرر وخيم؟ كيف يتم المزج بين المُحتوى المغلوط أو المضلّل وبين أنواع أُخرى من المُحتوى، مثل: المُحتوى الصحيح والمُحتوى الذي يُعزز الكراهية أو المُحتوى الترفيهي أو الآراء الشخصية؟ كيف يتم استغلال عالم المجهول من خلال تكتيكات التّضليل؟ هل أنّ هذه الرسائل تسعى إلى صرف الانتباه أو زعزعة مصداقية المُحتوى الصّادق و/أو الجهات الفاعلة المنخرطة في البحث عن الحقيقة (على غِرار الصحفيين والباحثين)؟

4. الوسائط:

- ما هي المواقع/الخدمات عبر الإنترنت ووسائل الإعلام الإخبارية التي تنتشر عليها المعلومات المُضلّلة؟ إلى أي مدى تقفز هذه المعلومات عبر هذه الأدوات الوسيطة، بحيث يتم نشرها أولاً على «الويب المظلم» وينتهي بها الأمر في وسائل الإعلام العادية؟

- كيف تنتشر هذه المعلومات؟ ما هو نموذج أعمال والخصائص الخوارزمية والسياسية للموقع أو التطبيق/الشبكة التي يتم استغلالها؟ هل تسعى الاستجابات إلى مُعالجة التحيُّز الخوارزمي الذي قد يؤدي إلى تفضيل المعلومات المُضلّلة؟ هل تعترف الاستجابات بأن « ... حُريّة التّعبير لا تعني حُريّة الانتشار» لأن «ليس هناك حق في التضخيم الخوارزمي» ولأنه لا ينبغي المساواة بين تنظيم المُحتوى الذي قد يتضمن الحد من التضخيم وزوال حُريّة التّعبير على الإنترنت (Diresta, 2018)؟ كذلك، هل هناك أدلة على وجود سلوك مُعد له (بما في ذلك السلوك المُفتعل) يقوم باستغلال نقاط الضعف، ليصوّر مُحتوى معيناً على أنه يتمتع بشعبية كبيرة (أو حتى بانتشار واسع) في حين أنه في الواقع قد يكون قد حقق هذا الانتشار من خلال اللعب المتعمد على الخوارزميات.
- هل تتصرف الوسائط بطرق تتسم بقدر كافٍ من المسؤولية والشفافية وينفذون الإجراءات اللازمة والمتناسبة للحد من انتشار المعلومات المُضلّلة؟

5. الجهات المستهدفة/المفسرة:

- من يتأثر بالمعلومات المُضلّلة؟ هل أنّ المستهدفين هم أفراد على غرار الصحفيين والعلماء؛ أم أنظمة (كالعمليات الانتخابية، الصحة العامة، المعايير الدولية)؛ أم مجتمعات المحلية؛ أم مؤسسات (مثل مراكز البحوث)؛ أم منظمات (بما في ذلك وسائل الإعلام الإخبارية)؛
- ما هي الاستجابات التي يتبعها هؤلاء على الإنترنت و/أو الإجراءات التي يتخذونها في الواقع خارج إطار الإنترنت؟ ويشمل هذا السؤال استجابات مختلفة على غرار عدم اتخاذ أي إجراء أو مشاركة المعلومات مما يؤدي إلى تأكيد محتواها أو الإعجاب بالمعلومات أو مشاركة المعلومات المُضلّلة بهدف كشف زيفها. هل من تقارير إخبارية غير نقدية (قد تؤدي إلى تحويل دور الصحفي/المنظمة الإخبارية المتواطئة من جهة مستهدفة بالمعلومات إلى وكيل تضليل)؟
- هل هناك تأثيرات على أرض الواقع لهذه الأعمال؟ على سبيل المثال، التأثير على التصويت، تشجيع الاحتجاجات، التحريض على جرائم الكراهية، مهاجمة الصحفيين، تقديم المشورة الطبية الخطيرة أو الخاطئة مما يثير مسألة ما إذا كانت الاستجابات تطال السياق الأوسع أو تقتصر على المُحتوى الموجود على الإنترنت.

باستخدام إطار «**IAMIT**» المختلط هذا كنقطة انطلاق لوضع تصور مفهومي للمعلومات المُضلّلة، من الممكن تصنيف الاستجابات على هذا الأساس. على وجه الخصوص، يمكننا التمييز بين التالي:

- الاستجابات التي تستهدف المُحرّضين على حملات التّضليل ووكلائها (الفصول 5.1 و5.2 و5.3).
- الاستجابات التي تستهدف تحديد المعلومات المُضلّلة، أي التحقُّق من الرسائل من حيث الزيف، وكشف المُحرّضين والوكلاء. (الفصول 4.1 و4.2)
- الاستجابات التي تهدف إلى الحد من إنتاج وتوزيع المعلومات المُضلّلة والسلوكيات ذات الصلة التي تنفذها الوسائط على وجه الخصوص (الفصول 6.1 و6.2 و6.3).
- الاستجابات التي تهدف إلى دعم الجهات المستهدفة/المفسرة لحملات تضليل (الفصول 7.1 و7.2 و7.3).

1.3.1 المعلومات المُضلّلة والدعاية

إنّ المعلومـات المُضلّـلة، كمـا تـم اسـتعراضها أعـلاه، تتميـز عـن مفهـوم الدعايـة وتتداخـل معـه. وإن فهـم القصـد أو النيـة أساسـي لفهـم الدعايـة، حيـث أن هـذه الأخيـرة تنطـوي علـى حملـة مدبـرة ومنظمـة. ولكـن ليـس هـذا دائمـاً حـال المعلومـات المُضلّـلة وفقـاً للتصـور المفهومـي الواسـع المعتمـد فـي هـذه الدراسـة.

وفـي الوقـت نفسـه، وكمـا هـو مذكـور فـي تعريـف قامـوس أوكسـفورد OED أعـلاه، وفـي الإعـلان المشـترك بشـأن حُريّة التّعبيـر و«الأخبـار الزائفـة» والمعلومـات المُضلّـلة والدعايـة[14]؛ قـد تتداخـل المعلومـات المُضلّـلة مـع الدعايـة التـي:

> **تعرّف بشكل محايد على أنها شكل منهجي من الإقناع الهادف الذي يحاول التأثير على عواطف ومواقف وآراء وأفعال جماهير مستهدفة معيّنة لأغراض أيديولوجية أو سياسية أو تجارية عن طريق النقل المضبوط للرسائل الأحادية الجانب (التي قد تكون أو لا تكون مبنيةً على وقائع) عبر قنوات ووسائل الإعلام الجماعية والمباشرة.**
> **(Nelson, 1996: p232-233)**

وهنـاك تاريـخ طويـل مـن التشـابك بيـن الدعايـة والمعلومـات المُضلّـلة (Posetti & Matthews, 2018). وتُسـتخدم تقنيـات الدعايـة المخادعـة أو «المظلمـة» علـى نطـاق واسـع (علـى سـبيل المثـال: الاسـتخدام الانتقائـي للحقائـق، الإقنـاع غيـر العـادل، إثـارة الخـوف)، علـى سـبيل المثـال فـي الحمـلات المناهضـة للاتحـاد الأوروبي، وسياسـات مـا بعـد الحقيقـة (Keane, 2018). والمواقـع الإلكترونيـة القائمـة علـى الأيديولوجيـة (علـى سـبيل المثـال: كراهيـة النسـاء أو كراهيـة الديـن الإسـلامي)، ووسـائل الإعـلام المتحزبـة بشـكل كبيـر. وغالبـاً مـا يكـون ذلـك بقصـد إحـداث تغييـرات سـلوكية فعليـة، علـى سـبيل المثـال: تعميـق الانقسـام الاجتماعـي، وزيـادة الاسـتقطاب، والتأثيـر علـى الـرأي العـام، أو التأثيـر علـى النتائـج السياسـية الرئيسـية.

فـي حيـن أنّ الدعايـة عـادةً مـا تُركّـز علـى الإقنـاع الاسـتراتيجي المُوجَّـه نحـو العمـل مـن خـلال تسـخير السـرديّات والهُويّـة والعاطفـة، فـإنّ «العمـل» الفكـري للمعلومـات المُضلّـلة (وفقـاً للمفهـوم المعتمـد هنـا) هـو فـي المقـام الأول «العبـث» بالحقائـق والمعرفـة (لا اسـتهداف المواقـف أو السـلوكيات). وعـادةً مـا تُسـخّر الدعايـة المعلومـات المُضلّـلة لتعزيـز غرضهـا الأكبـر. ومـع ذلـك، ففـي حيـن أنّ المعلومـات المُضلّـلة يمكـن أن تسـهم إسـهاماً كبيـراً فـي أيّ حزمـة دعائيـة، فـإنّ هـذه ليسـت تدخـلات مختلفـة مـن الناحيـة التحليليـة فحسـب، بـل قـد تكـون كل منهـا قائمـةً بذاتهـا. مـا يُعقّـد التقييـم هـو عندمـا تُدمـج المعلومـات المُضلّـلة بتقنيـات الدعايـة حـول الاختلافـات اللُّغويـة والثقافيـة والوطنيـة، لخلـق حواجـز اجتماعيـة جديـدة وتوسـيع دوائـر الانقسـام. وتُعـد تقنيّـة الدمـج هـذه سـمةً بـارزةً فـي الحمـلات السياسـية المثيـرة للانقسـام، سـواءً أكانـت تتـم داخـل الدولـة (مثـل: الحمـلات ذات الأهـداف القوميـة)، أو علـى يـد جهـات فاعلـة أجنبيـة (تهـدف علـى سـبيل المثـال إلـى زعزعـة الحيـاة السياسـية فـي دولـة أخـرى).

14 https://dictionary.cambridge.org/dictionary/english/misinformation

والأساس المنطقي للدمج بين تقنيات الدعاية وحملات التّضليل هو تعزيز مصداقية الرسالة. وهنا لابد من التأكيد على أنّ مصداقية الرسالة غير صحّتها، لأن الأولى تتعلق بالإدراك الذاتي لما إذا كانت معلومة معيّنة تتمتّع بالمصداقية، في حين أنّ صحة المعلومة تعتمد على التحقُّق وهو تقييم مستقل قائم على الأدلة.

بالإضافة إلى ذلك، يمكن أن يكون دمج تقنيات الدعاية والتّضليل استراتيجية للابتعاد عن استخدام مُحتوى كاذب بشكل واضح، واستخدام مُحتوى مأخوذ في غير سياقه يعتمد على التلاعب ويكون ذو طبيعة مضللة بهدف تشويه النظام البيئي للمعلومات.

2.3.1 المعلومات المُضلّلة وخطاب الكراهية

يعتمد خطاب الكراهية على التعامل مع جماعات معيّنة على أنها من «الآخرين»، وقد يستخدم المعلومات المُضلّلة كجزء من ترسانته لتدعيم رسالته مثلاً بمعلومات كاذبة وتعميمات عن فئة معينة من الأشخاص. وقد يكون هذا الخطاب جزءً من مبادرة دعائية، ولكن ليس بالضرورة.

وهنا لابد من إقامة تمييز هام بين المعلومات المُضلّلة من جهة وخطاب الكراهية من جهة أخرى، حيث يُفهم خطاب الكراهية على أنه «أيّ نوع من التواصل في الكلام أو الكتابة أو السلوك، يقوم بمهاجمة شخص أو جماعة ما أو يستخدم لغةً تحقيرية أو تمييزيةً ضدها على أساس الهُويّة أي بعبارات أُخرى على أساس الدين أو الانتماء الإثني أو الجنسية أو العرق أو اللون أو الأصل أو النوع الاجتماعي أو غير ذلك من عوامل الهُويّة» (UN, 2019; see also UNESCO, 2016). وغالباً ما تتقاطع الظاهرتان؛ على سبيل المثال عندما يتم استخدام الإساءة عبر الإنترنت والتّضليل جنباً إلى جنب، كما في حملات التشهير السياسي أو مهاجمة النساء الصحفيات بسبب كراهية النساء. ومع ذلك فكلا الظاهرتين مختلفتان من الناحية المفهومية، لأن المعلومات الكاذبة يمكن أن تكون قائمةً بذاتها من دون أن تمس الكراهية، على سبيل المثال في المعلومات المُضلّلة المعارضة للتلقيح ضد الفيروسات والأوبئة. أما خطاب الكراهية، من جانبه، فلا ينطوي دائماً وبالضرورة على معلومات مضللة؛ فقد يعتمد ببساطة على التّعبير عن الرأي والتحريض من دون استخدام الأكاذيب، بل عبر تضخيم المخاوف وكراهية الأجانب والنساء وغير ذلك من الأحكام المُسبقة.

وتركّز هذه الدراسة بوجه خاص على الاستجابات الهادفة إلى التصدي للمعلومات المُضلّلة التي يتم تقييمها من زاوية حُريّة التّعبير. ولذلك، لن يتم التطرق إلى الاستجابات التي تركز على خطاب الكراهية بشكل بحت. ومع ذلك، عندما تكون الاستجابات للمعلومات المُضلّلة مرتبطةً بخطاب الكراهية؛ سيجري بحث هذه الظاهرة من هذا المنظور.

4.1 المعلومات المُضلّلة، وحُريّة التّعبير، وأهداف الأمم المتحدة للتنمية المستدامة

فـي عـام 2015، حـددت الجمعيـة العامـة للأمـم المتحـدة سـبعة عشـر هدفـاً عالميـاً للتنميـة المُسـتدامة[15]. ويتأثـر عـدد منهـا باسـتخدام تكنولوجيـا النطـاق الواسـع وخدمـات شـركات التواصـل عبـر الإنترنـت لنشـر المعلومـات المُضلّـلة عبـر الإنترنـت؛ ولكـن أيضـاً لتنفيـذ بعـض الاسـتجابات للمعلومـات المُضلّـلة. وهــي:

- **الهدف 16 بشأن المجتمعات المُسالمة والشاملة للجميع والهدف الخامس من أهداف التنمية المستدامة بشأن المُساواة بين الجنسين:**

  - غالبـاً مـا تُسـتخدم المعلومـات المُضلّـلة علـى الإنترنـت لاسـتهداف الأفـراد (مثـل: السياسـيين والصحفييـن والمدافعيـن عـن حقـوق الإنسـان) والحكومـات والمجموعـات كالأقليـات العِرقيّـة والنسـاء والمجتمعـات القائمـة علـى الهُويّـة الجنسـانية والجماعـات الدينيـة والهويـات الدينيـة، بمـا فـي ذلـك مـن خـلال رسـائل قـد تُفضـي إلـى العنـف والكراهيـة والتمييـز.

  - لقـد تبيّـن أنّ الخوارزميـات التـي تسـتخدمها وسـائل التواصـل الاجتماعـي ومحـركات البحـث لتحديـد الأولويـات والتوصيـة بالمُحتوى (بمـا فـي ذلـك المعلومـات المُضلّلة) تقوم بإبـراز أولويات والتوصيـة بمحتويـات تُركّـز علـى لفـت النظـر وجـذب المُشـاركة، وهـي عُرضـة للتحيُّـز، وبالتالـي يُحتمـل أن تعمـل ضـد الشُّـموليّة. (مقـرر الأمـم المتحـدة الخـاص المعنـي بتعزيـز وحمايـة الحـق فـي حُريّـة الـرأي والتّعبيـر (2018أ)

- وتكتسي الغاية 16.1 من أهداف التنمية المستدامة أهميةً خاصةً في هذا التقرير الذي يتمحور حول حُريّة التّعبير والتّضليل لتركيزها على حصول الجمهور على المعلومات وحماية الحريات الأساسية.

  - تتعرّض حقوق المواطنين في التّعبير عن أنفسهم بحُريّة، والمشاركة على أساس مستنير في المناقشات المجتمعية عبر الإنترنت للخطر بسبب المعلومات المُضلّلة على الإنترنت، لا سيما عند انتشارها على نطاق واسع. ويمكن للمُحتوى الكاذب أن يُزعزع إيمان المواطنين وثقتهم في الحقائق والعلوم والعقلانية، فيثير بالتالي شكوكهم في المعلومات الموجودة على الإنترنت التي تتعارض مع آرائهم الشخصية. ومن شأن ذلك أن يثيهم عن المشاركة العامة، كما ويؤثر سلباً على مُمارستهم للحقوق والواجبات المتعلقة بالعمل المدني. ويكتسي ذلك أهميةً خاصةً بالنسبة إلى المواطنين والمجتمعات المحلية المستهدفة في حملات التّضليل التي تستخدم خطاب الكراهية لتغذية الانقسام وإشعال فتيل التوترات والقلاقل.

  - ومن الجدير بالذكر أن الحكومات والسياسيين هم من بين المُحرّضين الرئيسيين والجهات الأساسية الناقلة للمعلومات المُضلّلة (Bennen et al., 2020, Bradshaw & Howard 2019).

[15] https://sustainabledevelopment.un.org/?menu=1300

- يمكن أن يؤدي الاستخدام المتزايد لخوارزميات الذكاء الاصطناعي في انتقائية المعلومات المُضلّلة بشكل تلقائي (وأنواع المُحتوى الأخرى) إلى الإفراط في الرقابة على المُحتوى المشروع، وبالتالي انتهاك حُريّة المؤلف في التّعبير وحقه في الوصول إلى المعلومات. وقد تكون لهذه الخوارزميات أيضاً تحيزات كامنة وكما وقد تكون عرضةً للتلاعب.

- حملات التّضليل المنظمة والعضوية التي تستهدف الصحفيين (لا سيما الصحفيات) والوسائل الإخبارية كوسيلة لزعزعة ثقة المواطنين بالصحفيين وبالصحافة كمصادر موثوقة ومستقلة للمعلومات.

- مثال آخر هو الاستجابات القانونية غير المتناسبة للتضليل التي قد تؤدي في بعض الأحيان إلى إغلاق مواقع على الإنترنت والرقابة عليها، وتقييد الإبلاغ، وتجريم الصحافة، فضلاً عن التعاريف القانونية الغامضة للمعلومات المُضلّلة التي يمكن استخدامها لإسكات المعارضة السياسية أو الأصوات المعارضة (على سبيل المثال: عن طريق قوانين التصدي «للأخبار الزائفة»).

- **الهدف الرابع من أهداف التنمية المستدامة بشأن التعليم الجيّد الشّامل والمُنصف:**

- مع تزايد استخدام المواطنين لشبكة الإنترنت ومُحركات البحث للعثور على المعلومات لأغراض تعليمية، قد يؤثر وجود مستويات عالية من التّضليل على الإنترنت بشكل خطير على العمليات المعرفية الأساسية التي تضمن جودة التعليم، حيث أنّ العديد من المتعلمين يصبحون غير قادرين على الحكم على مصداقية المصادر وصدق المعلومات الموجودة على الإنترنت. وتزداد أهمية هذا الموضوع مع التحوّل الكبير إلى التعليم عن بُعد بفعل جائحة فيروس كورونا.

- ويمكن التّجربة على خوارزميات محركات البحث التي يستخدمها المواطنون للعثور على المعلومات لكي تعطي الأولوية للمعلومات المُضلّلة المنتشرة بشكل كبير، مما من شأنه أن يودي بالمتعلمين (وخاصةً الأطفال والأجيال الأكبر سناً) إلى تصديق نظريات المؤامرة وغيرها من المعلومات الخاطئة أو متدنية الجودة عبر الإنترنت.

- من ناحية إيجابية، توجد مشروعات صحافية استقصائية تركّز على مبادرات الدراية الإعلامية والمعلوماتية ومُبادرات التصدّي للمعلومات المُضلّلة، بما في ذلك الدراية بالبيانات؛ المُصممّة للاستجابة للمعلومات المُضلّلة الموجودة على الإنترنت والتي تهدف إلى التأثير بشكل إيجابي على ثقافة المواطنين ومعرفتهم وقدرتهم على التعرّف على المعلومات المُضلّلة وحمايتهم منها.

- **الهدف الثالث من أهداف التنمية المستدامة بشأن الحياة الصحية وتعزيز الرَّفاه - لجميع الأعمار:**

- إنّ المعلومات المُضلّلة المتعلقة بالصحة بشكلٍ عام -كما تبيّن خلال جائحة فيروس كورونا بما في ذلك الدعاية طويلة الأمد المضادة للّقاحات- تُعرّض حياة المواطنين الصحية ورفاههم للخطر (مثل: المعلومات المُضلّلة المتصلة بالحميات الغذائية). ونتيجةً للمعلومات المُضلّلة المضادة للتلقيح؛ شهدت معدلات تناول اللّقاحات انخفاضاً حاداً خلال السنوات

الأخيـرة (علـى سـبيل المثـال، فـي أفريقيـا (France24, 2020)، وآسـيـا (Power, 2020)، وأُوروبـا (Larson, 2018)، وأمريـكا الشـمالية (Burki, 2019).

والمعلومـات المُضلّلـة تتعـارض مـع أهـداف التنميـة المسـتدامة المتوافـق عليهـا. ومـع ذلـك، لا يـزال مـوردو هـذه المعلومـات (عـن قصـد أو عـن غيـر قصـد) وبفعـل مجموعـة مـن الدوافـع، يـرون مـن المفيـد أن يسـتثمروا الوقـت والمـوارد فـي إنتاجهـا وتعميمهـا– مسـتفيدين مـن نمـاذج الأعمـال والتكنولوجيـات التـي تعتمدهـا شـركات التواصـل عبـر الإنترنـت ووسـائل الإعـلام الإخباريـة للقيـام بذلـك.

"في الوقت نفسه، إنّ المعلومات المُضلّلة «لُعبة» لا فائز فيها على المدى الطويل. ولا شك في أن تصعيد حجم المعلومات المُضلّلة يؤدي في نهاية المطاف إلى التقليل من قيمة الحقائق بالنسبة إلى الجميع ويضع البشرية على طريق نحو انتشار الجهل في كل مكان. ونتيجةً لذلك أصبحت الإنجازات التي حقّقتها الحضارة حتى يومنا هذا والقائمة على حُريّة التّعبير مُعرّضةً للخطر. وعلى المحك؛ قضايا تتعلّق بالصحة والديمقراطية والأمن المالي والبيئة والحل السّلمي للصراعات والنزاعات والتوترات الاجتماعية والتماسك الاجتماعي، وأكثر من ذلك."

إن التّضليـل هـو ظاهـرة يصعـب علـى أيّ دولـة أو شـركة أن تتصـدى لهـا بمفردهـا –فهـي تتطلّـب التعـاون مـع الباحثيـن والمجتمـع المدنـي ووسـائل الإعـلام الإخباريـة. إنّ الاسـتهداف المدفـوع والدقيـق للأفـراد بمعلومـات مضللـة هـو مـن الأمثلـة التـي تدعـو إلـى تعـاون غيـر مسـبوق؛ واسـتخدام النّـد للنّـد لشـبكات الرسـائل الاجتماعيـة المغلقـة التـي تنشـر الأكاذيـب؛ هـو مـن قبيـل الأمثلـة الأخـرى المُدلّلـة علـى ذلـك.

لذلـك تسـعى هـذه الدراسـة إلـى النظـر فـي مجموعـة مـن الاسـتجابات التـي مـن شـأنها أن تمنـع المعلومـات المُضلّلـة وتحـد منهـا بـل وتتصـدى لهـا. وتقيّـم الفصـول التاليـة كافـة الإمكانيـات ونقـاط القـوة والضعـف فـي كلّ منهـا، فضـلاً عـن المخاطـر المحتملـة التـي قـد تشـكّلها علـى حُريّـة التّعبيـر، حيـث تسـعى جهـات فاعلـة متعـددة إلـى التصـدي للمعلومـات الخاطئـة.

وعلـى ذلـك، يعـرض الفصـل التالـي –الفصـل الثانـي– تصنيـف الاسـتجابات الهادفـة إلـى التصـدي للمعلومـات المُضلّلـة التـي تُشـكّل العمـود الفقـري لهـذه الدراسـة. ويقـدم الفصـل الثالـث مسـحاً مُفصّـلاً للبحـوث القائمـة، ويُسـلّط الضـوء علـى الفجـوات المعرفيّـة والفـرص المُتاحـة لمزيـد مـن الدراسـة. بعـد ذلـك، يتـم تحليـل كل نـوع مـن أنـواع الاسـتجابة الإحـدى عشـرة المعروضـة فـي التصنيـف الأصلـي الـذي تـم وضعـه لهـذه الدراسـة بشـكل منهجـي.

5.1 المنهجية

إنّ النتائج المعروضة هنا هي نتيجة بحث مكتبي أُجري في (أيلول/سبتمبر 2019– تموز/يوليو 2020) على يد فريق متعدد التخصُّصات يضم مؤلفين دوليين عملوا بطريقة تتسم بمُستوىً فائق من التعاون.

وسعى البحث الذي أُعد لغايات هذه الدراسة إلى إدراج مصادر تتعلّق بالبلدان في جميع القارات، بما في ذلك، حيثما أمكن (وفقاً للقُدرات اللُّغوية للباحثين)، مواد بلغات أُخرى غير الإنجليزية. كذلك ركّز الباحثون والباحثات على مكتبات وقواعد بيانات المؤسسات الأكاديمية والدُّول ومُنظّمات المُجتمع المدني ومواقع وسائل الإعلام الإخباريّة. وقد تمَّ الآن وضع العديد من هذه المصادر المُجمَّعة في لائحة المصادر الخاصّة بالدراسة.

كما وقام فريق الإشراف المؤلّف من خبراء بتقديم التعليقات على الدراسة. ويتألّف الفريق من: البروفيسور «فابريسيو بينيفينتو»، من جامعة ميناس جيرايس الاتحادية؛ والبروفيسور «ديفينا فراو-ميغز»، من جامعة السُّوربون الجديدة باريس 3؛ والبروفيسور «شيريان جورج»، من جامعة هونغ كونغ المعمدانية؛ والدكتورة «كلير واردل»، الرئيس التنفيذي لفيرست درافت First Draft؛ والبروفيسور «هيرمان واسرمان»، من جامعة كيب تاون. كما وعمل فريق البحث بشكل وثيق مع أمانة اليونسكو لرسم ملامح هذه الدراسة.



الفصل الثاني: تصنيف الاستجابات للمعلومات المُضلّلة

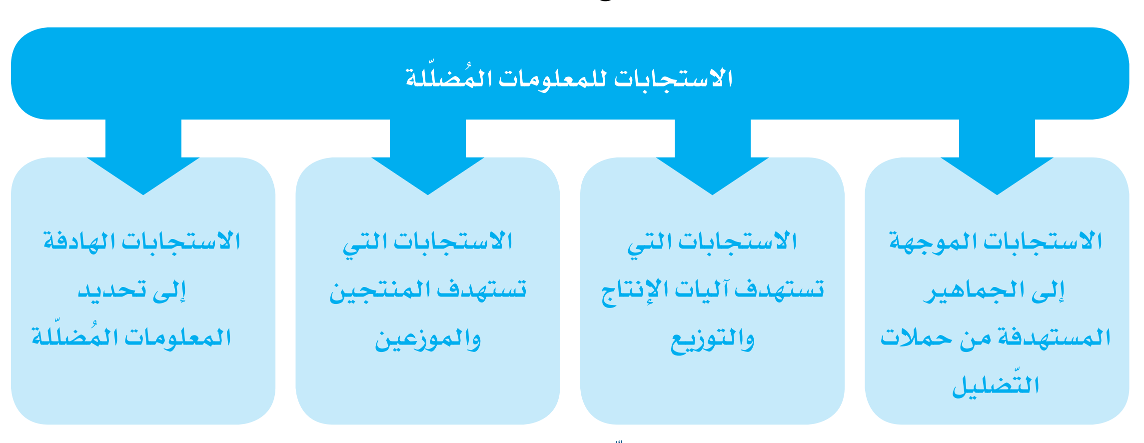
المُؤلّفون: كالينا بونتشيفا، جولي بوزيتي، دينيس تيسو، تريشا ماير، سام غريغوري، كلارا هنوت، ديانا ماينارد

يعرّف هذا الفصل بالتصنيف الهرمي للاستجابات المتعلقة بالمعلومات المُضلّلة والذي صيغ كجزء من البحث الذي أُجري لغايات هذا التقرير البحثي.

ووفقاً لهذا التصنيف، تُصَنّف الاستجابات للمعلومات المُضلّلة وفقاً لاستهدافها لجوانب معينة من الأزمة، وليس من حيث الجهات الفاعلة التي تقف وراءها (على غِرار شركات التواصل عبر الإنترنت، والحكومات، والمجتمع المدني، وما إلى ذلك). ويتيح وضع هذا الإطار تحديد المجموعة الكاملة من الجهات الفاعلة المشاركة في كل فئة من فئات الاستجابة وعبر هذه الفئات. على سبيل المثال، على الرغم من أن العديد من الجهات الفاعلة تميل في الوقت الحاضر إلى العمل بشكل مستقل أو أحادي؛ فمن شأن هذا التصنيف القائم على الاستجابة أن يفتح المجال أمام تآزر محتمل في المستقبل يعمل على بلورة مقاربة لتوفير هذه الاستجابات تضم مختلف أصحاب المصلحة داخل فئات التدخل هذه وفي ما بينها.

ويتمثل الدافع الرئيسي الثاني وراء وضع هذا التصنيف القائم على الاستجابة في أنه يسمح بإجراء تحليل لأثر كل نوع من أنواع الاستجابة على حُريّة التّعبير (وعند الاقتضاء، على حقوق الإنسان الأساسية الأخرى كاحترام الخصوصية). ويتم بشكل خاص تقييم كل فئة من فئات الاستجابة ليس فقط من حيث مَواطن قوتها وضعفها بشكل عام ولكن وعلى وجه التحديد في ما يتعلق بحُريّة التّعبير.

ويميّز تصنيف الاستجابات للمعلومات المُضلّلة أربع فئات أساسية (رجاءً مراجعة الشكل 1 أدناه):



الشكل 1. الاستجابات الرئيسية للمعلومات المُضلّلة

إنّ الفئات الواردة في هذا التصنيف لا تنفي بعضها البعض. أي أن هناك بعض التدخلات التي تنتمي إلى أكثر من فئة من التصنيف، حتى لو كانت هناك أبعاد تشمل فئات أخرى على سبيل المثال. وفي هذه الحالة؛ فإنها توضع في إطار إحدى الفئات ولكن يتم الإحالة إليها في فصول أخرى متى كان ذلك مناسباً. وعلى سبيل المثال، فإنّ مبادرات التحقُّق من الوقائع الخاصة بالانتخابات ذات صلة بالفصل الذي يناقش الاستجابات الانتخابية (5.3) وبالفصل المتعلق باستجابات التحقُّق من الوقائع (4.1)، ولذلك يتم تناولها في المقام الأول في الفصل 5.3، ولكن تتم الإحالة إليها أيضاً في الفصل 4.1.

تصنيف الاستجابات للمعلومات المُضلّلة:

الاستجابات للمعلومات المُضلّلة

الاستجابات الهادفة إلى تحديد المعلومات المُضلّلة:
- استجابات الرصد والتحقُّق من الوقائع
- الاستجابات الاستقصائية

الاستجابات التي تستهدف المنتجين والموزعين:
- الاستجابات التشريعية وما قبل التشريعية والسياساتية
- الحملات الوطنية والدولية المناهضة للمعلومات المُضلّلة
- الاستجابات الانتخابية

الاستجابات التي تستهدف آليات الإنتاج والتوزيع:
- الاستجابات المتعلقة بتنظيم المُحتوى
- الاستجابات التقنية والخوارزمية
- استجابات حجب إمكانية جني الأرباح

الموجهة إلى الجماهير المستهدفة من المعلومات المُضلّلة:
- الاستجابات الأخلاقية والمعيارية
- الاستجابات التثقيفية
- استجابات التمكين وتحديد المصداقية

الشكل 2. الاستجابات الأربع الرئيسية والاستجابات الفرعية الإحدى عشرة.

بمزيد من التفصيل، تتضمّن الاستجابات للمعلومات المُضلّلة رصد 0 وتحليل قنوات مشاركة ونشر المعلومات (مثل: وسائل التواصل الاجتماعي وتطبيقات المراسلة، ووسائل الإعلام الإخبارية، والمواقع الإلكترونية) لتحديد وجود معلومات مُضلّلة. والهدف هنا هو تحديد وجود المعلومات المُضلّلة ومدى التّضليل. على وجه الخصوص، تم تمييز فئتين فرعيتين من الاستجابات لتحديد المعلومات المُضلّلة:

- **استجابات الرصد والتحقُّق من الوقائع** – التي تُنفذ معظمها على يد الوكالات الإخبارية وشركات التواصل عبر الإنترنت والأوساط الأكاديمية ومنظمات المجتمع المدني والمنظمات المستقلة العاملة في التحقُّق من الوقائع، وكذلك (حيثما وُجدت الإمكانية) الشراكات بين مثل هذه المنظمات والكيانات.

- **الاستجابات الاستقصائية**، التي تتخطى مسألة معرفة ما إذا كانت رسالة معيّنة/مُحتوى معين غير صحيح أم لا(جزئياً)، لتوفر معلومات متعمقة حول حملات التّضليل، بما في ذلك الجهات الفاعلة الأصلية ودرجة الانتشار والمجتمعات المتأثرة.

أما الفئة الرئيسية الثانية فهي الاستجابات الموجهة إلى منتجي المعلومات الخاطئة وموزعيها عن طريق تغيير البيئة التي تحكم سلوكهم وتُشكّله (الاستجابات القانونية والسياستية):

- الاستجابات التشريعية وما قبل التشريعية والسياستية، التي تشمل تدخلات ناظمة للتصدي لمعلومات المُضلّلة.
- الحملات الوطنية والدولية المناهضة للمعلومات المُضلّلة، والتي تميل إلى التركيز على بناء سرديات مضادة.
- الاستجابات الانتخابية المصممة خصيصاً للكشف عن المعلومات المُضلّلة التي تنتشر خلال الانتخابات؛ فضلاً عن تعقبها ومكافحتها. وعلى الرغم من أنّ للمعلومات المُضلّلة عبر الإنترنت تستهدف أموراً مهمةً أخرى (مثل: التلقيح ضد الفيروسات وغيرها من المعلومات الصحية)، تم إدراج فئة خاصة بالاستجابات التي تستهدف المعلومات المُضلّلة الانتخابية وذلك بسبب تأثيرها على العمليات الديمقراطية وحقوق المواطنين. وعادةً ما تنطوي هذه الفئة من الاستجابات، بطبيعتها، على مزيج من استجابات الرصد والتحقُّق من الوقائع والاستجابات القانونية والتنظيمية والتقنية وغير ذلك؛ وسيجري الإحالة إليها حسب الاقتضاء. ويُسلّط هذا الضوء على المقاربة متعددة الأبعاد المطلوبة لمُكافحة المعلومات المُضلّلة المتصلة بالانتخابات لانطوائها على خطر الإضرار بالمؤسسات الديمقراطية.

أما الفئة الرئيسية الثالثة، فهي استجابات تستهدف عمليات إنتاج ونشر المعلومات المُضلّلة، تشمل الاستجابات المتعلقة بتنظيم المُحتوى، وحجب الربح المالي، ووضع المعلومات في إطارها الأصلي واستخدام تقنية التشغيل الآلي:

- وتتناول الاستجابات التي تُعنى بتنظيم المُحتوى في المقام الأول سياسة التحرير وتطوير المُحتوى و«المعايير المجتمعية» على الرغم من أنّه قد يكون لبعضها بعد تكنولوجي أيضاً؛ وسيتم الإحالة إلى ذلك عند الاقتضاء.
- الاستجابات التقنية والخوارزمية التي تستخدم الخوارزميات و/أو الذكاء الاصطناعي (AI) من أجل الكشف عن المعلومات المُضلّلة والحد منها، أو توفير سياق أو معلومات إضافية عن عناصر ومنشورات فردية. ويمكن تطبيقها من خلال المنصات الاجتماعية ومحركات مشاركة الفيديو ومحركات البحث نفسها، كما ويمكن أن تُشكّل أيضاً أدوات مستخدمة من طرف ثالث (مثل: البرامج المضافة للمتصفح plug-ins) أو طرق تجريبية صادرة عن البحوث الأكاديمية.
- الاستجابات الهادفة إلى حجب الربح المالي وهي مصممة لوقف تحقيق الدخل والربح من المعلومات المُضلّلة وبالتالي منع التحفيز على وضع عناوين خاطفة clickbait أو مواقع الأخبار الزائفة وأنواع أخرى من التّضليل الهادف إلى تحقيق الربح.

أمـا الفئـة الجامعـة الرابعـة، فتجمـع اسـتجابات راميـة إلـى دعـم الجماهيـر المسـتهدفة مـن حمـلات التّضليـل (أي «الضحايـا» المحتمليـن للمعلومـات المُضلّلـة). وتشـمل هـذه الاسـتجابات مبـادرات تسـعى إلـى وضـع مبـادئ توجيهيـة، وتوصيـات، وقـرارات، وتعزيـز الدرايـة الاعلاميـة والمعلوماتيـة، ومبـادرات تسـعى إلـى تحديـد مصداقيـة المُحتـوى. وتصنـف هـذه الاسـتجابات المختلفـة فـي فئـات فرعيـة هـي:

- الاستجابات الأخلاقية والمعيارية المعتمدة على المستويات الدولية والإقليمية والمحلية والتي تنطوي على إدانة علنية لأعمال التّضليل أو وتوصيات وقرارات ترمي إلى إحباط هذه الأعمال وتوعية الجمهور بهذه المسائل.
- الاستجابات التثقيفية التي تهدف إلى تعزيز دراية المواطنين الإعلامية والمعلوماتية، والتفكير الناقد والتحقُّق في ما يخص استخدام المعلومات عبر الإنترنت، فضلاً عن تدريب الصحفيين.
- جهود التمكين وتحديد المصداقية التي تسعى إلى بناء أدوات التحقُّق من المُحتوى ومؤشرات حول المُحتوى على الإنترنت، وهي أدوات مساعدة بإمكانها مساعدة المواطنين والصحفيين على تجنب الوقوع فريسة للمعلومات المُضلّلة على الإنترنت. وقد تهدف هذه الجهود أيضاً إلى التأثير على تنظيم المُحتوى لفائدة بروز وتضخيم مُحتوىً معين – وقد تم تضمينها في الاستجابات الخاصة بتنظيم المُحتوى أعلاه.

بعـد اسـتعراض مفصـل للأدبيـات وإتمـام مسـح للمشـهد فـي الفصـل الثالـث، يقـوم هـذا التقريـر بتحديـد وتحليـل وتقييـم الاسـتجابات للمعلومـات المُضلّلـة وفقـاً لهـذا التصنيـف. فـي كل حالـة، يصـار إلـى تفصيـل الخصائـص المميـزة للفئـة وطـرح مجموعـة مشـتركة مـن الأسـئلة؛ بهـدف التحفيـز علـى تفسـير أُسُـس كل نـوع مـن أنـواع الاسـتجابة. وهـذه الأسـئلة هـي:

- من هي الجهات و/أو ما هي الأمور التي تسعى نوع الاستجابة إلى رصدها؟
- من هو الجمهور المستهدف من نوع الاستجابة/من تسعى إلى مُساعدته؟
- ما هي مخرجات هذا النوع من الاستجابات (مثل: المنشورات والقوانين)؟
- من هي الجهات الفاعلة التي تقف وراء هذه الاستجابات، ومن يُموّلها (حيثما كان ذلك معلوماً)؟
- كيف يتم تقييم فعالية هذه الاستجابات؟
- ما هي نظرية التغيير التي تعتمدها؟
- ما هي نقاط قوتها وضعفها بوجه عام وفي ما يتعلق بحُريّة التّعبير بوجه خاص؟
- ما هي الثغرات وأوجه التآزُر المحتملة التي تم تحديدها في سياق التحليل؟

وأخيـراً، عنـد الاقتضـاء، يتـم تنـاول أزمـة المعلومـات المُضلّلـة المتعلقـة بجائحـة فيـروس كورونـا (Posetti and Bontcheva 2020a; Posetti and Bontcheva 2020b) مـن خـلال دراسـة حالـة مصغـرة ضمـن الفصـول.



الفصل الثالث: إطار البحث والثّغرات

المُؤلّفون: ديانا ماينارد، جولي بوزيتي، كالّينا بونتشيفا، ودينيس تيسو

يضع هذا الفصل محور المعلومات المُضلّلة لهذا التقرير في سياق الأُطُر النظريّة القائمة والتقارير السّابقة بشأن هذا الموضوع. كما ويربط المعلومات المُضلّلة بحُريّة التّعبير وأهداف التنمية المستدامة ذات الصلة بلجنة النطاق الواسع (وتحديداً الهدف 16 المتعلق بالمجتمعات السلمية والشاملة، والغاية 16.10 بشأن سلاسة وصول الجمهور إلى المعلومات والتمتع بالحريات الأساسية). على وجه الخصوص، لا يُركّز هذا الفصل على المُحتوى غير الصحيح في حد ذاته فحسب، بل على الجهات الفاعلة وعلى الدوافع التي تؤدي بها إلى مشاركة المعلومات المُضلّلة كما وعلى أهداف حملات التّضليل، وبالتالي يناقش تضخيم هذا النوع من المُحتوى والتلاعب به. وبالإضافة إلى ذلك، يدرس الفصل الأدبيات المتعلقة بطرق الاستجابة للمعلومات المُضلّلة. ثم يناقش بمزيد من التعمُّق الثغرات في البحوث التي أُجريت قبل أوائل عام 2020، ويُفصّل الإسهامات الجديدة لهذه الدراسة، مقارنةً بالتقارير السابقة عن مظاهر التّضليل.

ومع تحليل الثغرات هذا، يتم التركيز بشكل خاص على تأثير المعلومات المُضلّلة على المجتمعات وعلى كيفية تلقّيها من قِبل الجمهور، وذلك من خلال مراجعة أدبيات العلوم المعرفية وعلم النفس الاجتماعي، بالإضافة إلى الأدبيات المتعلقة بالسياسة والصحافة وعلوم الإعلام والتواصل والقانون. ولا يشمل الاستعراض الأدبيات الأكاديمية فحسب، بل كذلك التقارير السياسية الصادرة عن المجموعات التي تمثل القطاعات والمجتمع المدني، والأوراق البيضاء، والكتب والمنشورات الموجهة إلى الجمهور العام، والمقالات الإخبارية والمجلات على الإنترنت. ولكن لابدّ من التأكيد على أن القصد من هذا الاستعراض ليس أن يكون شاملاً بل أن يذكر بعض الاتجاهات والثغرات الرئيسية في مجال البحوث، وتلك الموجودة في الاستجابات أيضاً.

ولا يسعى هذا الفصل إلى تقييم جودة الأعمال المختارة تقييماً حتمياً، بل إلى محاولة فهم طبيعة مجموعة من البحوث حول المعلومات المُضلّلة، وأُسسها النظرية، وأنواع الدراسات التي أُجريت، والطُّرق التي نوقشت بها المعلومات المُضلّلة في الأوساط الصحفية والأكاديمية والرسمية (أي الجهات الفاعلة الحكومية وغير الحكومية)، وكذلك كيف تم نشرها على نطاق أوسع على عامة الجمهور. ثم يقوم هذا الفصل بتلخيص النتائج، مع التركيز على بعض المجالات الرئيسية مثل: المعلومات المُضلّلة السياسية وصنع السياسات، ويُسلّط الضوء على بعض الاتجاهات الناشئة، ثم يقوم بمناقشة القيود المفروضة على هذه البحوث ووضع هذا التقرير ضمن الأدبيات الأكاديمية.

وبالتالي فإن هدف هذا الاستعراض هو فهم ما يمكن تعلّمه حول أنواع المعلومات المُضلّلة الموجودة؛ أي من يحرّض عليها ومن ينشرها ولماذا؛ والجهات التي تستهدفها هذه المعلومات؛ والفعالية النسبية لمختلف أنواع المعلومات المُضلّلة والطرق المُستخدَمة؛ والاستجابات القائمة للتصدي لها؛ والتوصيات للمستقبل لا سيما في ضوء المخاوف المتعلقة بحُريّة التّعبير. وهذا يمهّد للفصول التالية التي تبحث بمزيد من التعمُّق في مختلف الاستجابات للمعلومات المُضلّلة فضلاً عن نظريات التغيير المرتبطة بها، وفوائدها ومثالبها المحتملة و/أو المؤكدة.

ويتم التركيز على النظر في التهديدات الجديدة، على غرار تقويض حُريّة التّعبير عن طريق استخدام أساليب الانتقائية العشوائية باستخدام الذكاء الاصطناعي، وتزايُد المواد الإعلامية المُلفقة (وتُسمى أيضاً بـ«المواد الزائفة الفجّة») كأساليب جديدة للتضليل. وقد بدأت هذه المشكلة في الظهور فعلياً؛

حيث تم استهداف العديد من السياسيين والصحفيين بادعاءات حول قيامهم بسلوكيات جنسية غير ملائمة عبر نشر مقاطع فيديو لهم تم التلاعب بها و/أو تزييفها عبر برمجيات التزييف الفج/العميق[16].

وهناك اتجاه حديث تم التقليل من أهميته إلى حد كبير في الماضي، وهو صعود السرديات المُتناقضة (Decker, 2019)، حيث لا تشمل استراتيجيات التّضليل نشر نظريات المؤامرة البسيطة والأكاذيب الصريحة فحسب، بل ظواهر أكثر تعقيداً يتم فيها شحن المعلومات الحقيقية والخاطئة عاطفياً وربطها ببعضها البعض في شبكات معقدة مُصمَّمة خصيصاً لإرباك الأفراد وإشعارهم بالصّدمة وتشتيت انتباههم فضلاً عن إحباطهم وجعل الباحثين عن الحقيقة دائماً في وضعيةٍ دفاعية. وإذا كان توفير المعلومات شرطاً لتمكين الجمهور، فالتّضليل يعمل على بعثرة المعلومات وتشويه مصداقيتها؛ غالباً بهدف الإضعاف وإثارة الارتباك. ومن الأمثلة على ذلك: الممارسة المُسمّاة «زرع الشك» المعروفة باسم «gaslighting»، وهي استراتيجية قوية تهدف إلى السيطرة على الأفراد من خلال ممارسة السلطة والتلاعب بتصوراتهم للواقع– وبالتالي زرع الخوف والاضطراب، ومن ثم الظهور بثوب من يُقدّم الحلول (Keane, 2018). وبإمكان تقنيات التّضليل هذه والتي غالباً ما توصف بأنها «استعمال المعلومات كسلاح»، أن تُدمّر التماسك الاجتماعي وتُهدّد الديمقراطية (Hansen, 2017). وقد تُحفّز طلب الجمهور العام لمزيد من اليقين والسيطرة من قبل السياسيين، مما يُهدد بفرض مزيد من القيود على حُريّة التّعبير، ويعزز الاستبداد الاجتماعي (Flore et al., 2019). ومن ناحية أخرى، تتقدم الاستجابات للمعلومات المُضلّلة وتتطور وتسعى إلى إدماج معايير حقوق الإنسان للتصدي للأضرار المحتملة.

16 يُرجَى مراجعة الأمثلة من فنلندا (Aro, 2016) والهند (Ayyub, 2015) وجنوب أفريقيا (Haffajee & Davies, 2017) وغيرها.

1.3 الأُطر المفاهيمية التي تسمح بفهم المعلومات المُضلّلة المُعاصرة

لقد ظهرت في السنوات الأخيرة موجة من البحوث تسعى ليس فقط إلى دراسة طبيعة ونطاق المعلومات المُضلّلة بل أُسسها النفسية وأُطرها النظرية أيضاً. وتعكس هذه الأُطُر جوانب مختلفة. ولكلّ من (Derakhshan & Wardle, 2017) نظرةً شاملة للموضوع، حيث يعتبران أنّ «الاضطراب في المعلومات» مشكلة ثلاثية الأبعاد تضم (وفقاً لتعريفهما) «المعلومات المُضلّلة» و«المعلومات المغلوطة» و«المعلومات الضارة». وتعتمد وجهات النظر الأخرى أنظمة تصنيف أضيق على غِرار توصيف حملات التّضليل السياسية لمختبر البحث الرقمي الجنائي أي Research Lab Digital Forensic باللغة الإنكليزية (Brooking et al., 2020).

1.1.3 «اضطراب المعلومات» و«حرب المعلومات»

في تقريرهما لمجلس أوروبا، يفصّل (Wardle & Derakhshan, 2017) مفاهيمهما، ويقدمان ملخصاً للبحوث والتقارير والمبادرات العملية ذات الصلة التي تم إنتاجها حول هذا الموضوع حتى منتصف عام 2017. ويبحثان في التقرير الذي وضعاه في الأفكار والحلول التي لابد أن تُطبَّق عبر وسائل الإعلام الإخبارية ومنصات التواصل الاجتماعي كما في الأفكار والحلول التي لابد من تطبيقها لمصلحة وسائل الإعلام ومنصات التواصل. بالإضافة إلى ذلك، يقدمان دراسة للاتجاهات والتداعيات المستقبلية. ويشمل ذلك التركيز على استخدام الذكاء الاصطناعي، ليس فقط لاكتشاف المعلومات المُضلّلة بل أيضاً لإنتاجها. كما ويورد التقرير 34 توصيةً لشركات التكنولوجيا والحكومات والمنظمات الإعلامية والهيئات الممولة والمواطنين بشكل عام. ولا شك في أن العديد من هذه التوصيات هو أصلاً موضع تنفيذ بشكلٍ ما (على سبيل المثال، لقد بدأت بعض شركات التكنولوجيا بالفعل ببناء أدوات للتحقق من الوقائع) فيما بعض التوصيات الأخرى تحتاج إلى مزيد من التفسير (على سبيل المثال، كيف يمكن للمجتمع المدني أن «يكون وسيطاً نزيهاً»، أو كيف يمكن لوزارات التعليم أن «تعمل مع المكتبات»).

وقد بُني الإطار المفاهيمي لـ Wardle وDerakhshan على عملهما السابق المُعنون «الأخبار الزائفة.. موضوعٌ مُعقّد» (Wardle, 2017a) أو «Fake News, It's Complicated» الذي يحدد سبعة أنواع من المعلومات المغلوطة والمعلومات المُضلّلة، بدءً من التقليد السّاخر والتهكُّم (وهما نوعان قد يؤديان إلى ما يسمى بالمعلومات المغلوطة كونهما قابلين للتفسيرات الخاطئة) وصولاً إلى التلفيق المتعمد. ويصف الإطار عملية إنتاج المعلومات المغلوطة ونشرها بأنها عملية ثلاثية مكوّنة من وكيل ورسالة ومفسر (هدف). ومع ذلك، وكما ورد في مقدمة هذه الدراسة، فإن الطابع العملي لهذا الإطار يواجه تحدياً أساسياً يتمثل في أنّ التمييز بين المعلومات المغلوطة والمعلومات المُضلّلة قائم بشكل رئيسي على الدافع وإدراك الزيف. فالمشكلة هي أنّ الدوافع متنوعة ومتناقضة في كثير من الأحيان، كما وأنها غالباً غير واضحة أيضاً. إضافةً إلى ذلك، قد يؤدي هذا التمييز إلى المبالغة في التركيز على القصد على حساب الأثر المشترك للنوعين. على سبيل المثال، إذا قرر الأفراد رفض التلقيح كنتيجة لتعرضهم لمُحتوى غير صحيح حول الموضوع، فإنّ النتيجة هي نفسها سواء تمت مشاركة هذا المُحتوى عبر المعلومات المغلوطة أو المعلومات المُضلّلة. ولا تكتسي الدوافع أهميةً على الرغم من صعوبة تحديدها في كثير من الأحيان، إلا عند تقييم مدى مُلاءمة استجابة معينة لفائدة تحديدها للدوافع. لذلك، تُركّز هذه الدراسة على الاستجابات الاستقصائية كمصدر للمعرفة؛ من شأنه إرشاد أنواع الاستجابات الأُخرى.

ومـن الاعتبـارات الأخـرى فـي مـا يخـص إطـار «اضطـراب المعلومـات» إمكانيـة قيـام هـذا الإطـار بتفضيـل التمييـز الثنائـي بيـن المعلومـات «المضطربـة» والمعلومـات «غيـر المضطربـة»، وبالتالـي تعزيـز وجهـة النظـر التـي تختـزل صحـة المعلومـات بمقاربـة الأبيـض والأسـود، مـع إغفـال أو إنـكار نطـاق المجهـول سـواءً أكان هـذا المجهـول شـكوكاً علميـةً أو مناقشـات سياسـية مشـروعة. وهنـاك مسـألة أخـرى هـي أنّ «المعلومـات الضـارة» قـد تُفسَّـر بطريقـةٍ تـؤدي إلـى وصـم مجموعـة مـن السـرديات غيـر التّضليليـة التـي تقـوم بفعـل طبيعتهـا باختيـار وتفسيـر حقائـق وسـياقات معينـة كجـزء مـن مناقشـتها المشـروعة للمعنـى.

فـي ضـوء ذلـك، يعمـل هـذا البحـث علـى مسـتوى أكثـر تجريـداً؛ أبعـد مـن التمييـز بيـن المُحتـوى المغلـوط أو المُضلّـل وفقـاً لمعاييـر تتعلّـق بالدوافـع، فيركّـز بـدلاً مـن ذلـك علـى كافـة المحتويـات الخاطئـة التـي قـد تـؤدي إلـى ضـرر محـدد. وهـذا يوفّـر وسـيلةً لتقييـم كافـة الاسـتجابات فـي عملهـا علـى بنـاء إطارهـا المفهومـي.

وقـد قـام (Derakhshan 2019) بتقييـم مقاربـة لقضيـة المعلومـات المُضلّلـة تعتمـد تركيـزاً اسـتراتيجي وذلـك فـي تقريـره المُعنـون «Disinfo Wars» أو «حـروب التّضليـل». ويناقـش فيـه العلاقـة بيـن وكلاء المعلومـات المُضلّلـة وأهدافهـا فـي مـا يسـميه «تصنيـف حـرب المعلومـات». وبنـاءً علـى ذلـك، فـإن المقاربـة المعتمـدة تحصـر فكـرة المعلومـات المُضلّلـة بمفهـوم أضيـق بكثيـر يتمحـور حـول الاسـتراتيجية السياسـية بـل وحتـى العسـكرية. ومـن الأمثلـة علـى ذلـك منظـور «شـركة رانـد» حـول «حمـلات المعلومـات»/ «حمـلات التأثيـر» والتـي تربـط هـذه المُصطلحـات بـ«جمـع المعلومـات التكتيكيـة عـن الخصـم ونشـر الدعايـة سـعياً إلـى تحقيـق ميـزة تنافسـية عليـه»[17]. وقـد اعتمـد فريـق التواصـل الاسـتراتيجي للشـرق التابـع لدائـرة العمـل الخارجـي الأوروبيـة موقفـاً مماثـلاً[18]. ويجـادل «ديراخشـان» بـأن معظـم الأمـوال والجهـود التـي تُنفـق علـى مُكافحـة المعلومـات المُضلّلـة فـي «حـرب المعلومـات» ينبغـي أن تُركّـز علـى المسـتهدفين، أي الجهـات الفاعلـة غيـر الحكوميـة علـى غـرار وسـائل الإعـلام.

فـي حيـن أن حُجَّتـه تُغطّـي مجموعـةً واسـعةً مـن الأنشـطة، إلا أنهـا تركـز إلـى حـد مـا علـى المُحتـوى غيـر الصحيـح الـذي يتـم نشـره لدوافـع معينـة، كمـا فـي مـا كتبـه «واردل» سـابقاً (Wardle, 2017a). وكمـا ذُكـر أعـلاه، فـإن هـذا الأمـر معقـد مـن الناحيـة العمليـة، وهـو يتجـاوز حتـى المسـائل المعقـدة المتعلقـة بالإسـناد. وبالإضافـة إلـى ذلـك، وفـي حيـن أن التركيـز الاسـتراتيجي علـى الأبعـاد والخصائـص الجغرافيـة والسياسـية أمـر مهـم، فـإن المجتمـع يواجـه أيضـاً مسـألة المعلومـات المُضلّلـة باعتبارهـا مشـكلةً أوسـع بكثيـر. وهنـاك أيضـاً نقـص فـي الأدلـة لمعرفـة مـا إذا كان العمـل مـع مجموعـة معينـة (وسـائل الإعـلام، أو الجمهـور العـام) أقـل أو أكثـر فعاليـةً مـن العمـل مـع مجموعـة أخـرى.

....

17 https://www.rand.org/topics/information-operations.html

18 https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1006

2.1.3 حملات التّضليل السياسي

لقد اعتمد الباحثون في مختبر البحث الرقمي الجنائي منظوراً يركز بشكل أقل على الاستعارات الحربية ويتعامل مع المعلومات المُضلّلة السياسية كمفهوم أوسع. كما واقترحت «وحدة جيكسو» التابعة لغوغل أي Google's Jigsaw (وهي وحدة تعمل على مُكافحة «العواقب غير المقصودة» للتكنولوجيا الرقمية) واختبرت نظام تصنيف لحملات التّضليل السياسي يقوم على 150 خياراً متغيراً (Brooking et al., 2020). والهدف الرئيسي من هذا الإطار هو وضع وصف لأنواع مختلفة جداً من الجهود السياسية المُضلّلة ومقارنتها ببعضها البعض. ويحتوي الإطار على ست فئات رئيسية: الهدف، المنصة، المُحتوى، الطريقة، الإسناد، والقصد/النية. وتقسم كل فئة من هذه الفئات إلى فئات أخرى وإلى فئات فرعية أخرى. ويبيّن الجدول أدناه الفئات من المستويين الأول والثاني، مع بعض الأمثلة من المستوى الثالث. وعادةً ما تكون فئات المستوى الثالث ثنائيةً (على سبيل المثال، ما إذا كان الهدف مرتبطاً بالحكومة أم لا)، على الرغم من أن المقاييس الكمية تنطوي على أرقام أو نسب، وبعضها يعتمد استجابات حرة. بالإضافة إلى ذلك، تحتوي جميع فئات المستوى الثاني على فئة يمكن إضافة ملاحظات حرة فيها، وبعضها يحتوي أيضاً على فئة فرعية «أخرى».

المستوى الأول	المستوى الثاني	ملاحظات/أمثلة
الهدف	الهدف الأساسي	الحكومة، الأحزاب السياسية، قطاع الأعمال، الجماعات العرقية، الأفراد المؤثرون (بمن فيهم الصحفيون) ومجموعات الأفراد.. إلخ
	المقاييس الكمية	مؤشرات/تقييم الاستقرار السياسي، حُريّة الإنترنت، تعداد اللاجئين وما إلى ذلك.
	الأحداث المتزامنة	الحرب والانتخابات وما إلى ذلك
	الهدف الثانوي	الحرب والانتخابات وما إلى ذلك
	الهدف الثالث	الحرب والانتخابات وما إلى ذلك
المنصات	عالم الإنترنت المفتوح	وسائل الإعلام الحكومية، وسائل الإعلام المستقلة، وسائل الإعلام الأخرى
	وسائل التواصل الاجتماعي	فيس بوك، انستغرام، تويتر، المنتديات.. إلخ
	منصات المراسلة	واتساب، تلغرام، وي-تشات Wechat، والرسائل القصيرة.. إلخ
	الإعلانات	(يتم شراؤها من قبل الجهات المُضلّلة لنشر رسائل تضليلية بما في ذلك على وسائل التواصل الاجتماعي وعالم الإنترنت المفتوح)
	البريد الإلكتروني	
المُحتوى	اللُّغة	
	الموضوعات	ما هو مضمون المعلومات المُضلّلة، على سبيل المثال: الحكومة، الجيش، الانتخابات، الإرهاب، العنصرية، وما إلى ذلك
الطرق	التكتيكات	حشد اللّجان الإلكترونية، شبكات روبوتات الإنترنت، الشبكات الدُّمى، التلاعب بمحركات البحث، القرصنة، الصُّور الزائفة بفجاجة.. إلخ
	تقنيات السرد	– بنّاءة (على سبيل المثال: سلوك القطيع bandwagon، ودعاية شعبية زائفة astroturfing)؛ – مُدمرة (مثل: الترهيب والتشهير)؛ – مُنحرفة (الاستفزاز على الإنترنت trolling، والإغراق flooding)

الإسناد	المُضلّل الأساسي	بلد، كتلة أو غيره
	فئة المُضلّل	كما في الفئة المستهدفة، على سبيل المثال: الحكومة، الأحزاب السياسية، قطاع الأعمال، الشخصيات المؤثرة، الأقليات
	الكمية المقاييس	بالنسبة إلى هدف معين، على سبيل المثال: بيانات الاستقرار السياسي، وحُريّة الإنترنت، تعداد اللاجئين
	الأحداث المتزامنة	بالنسبة إلى الأهداف مثل: الحرب والانتخابات وما إلى ذلك
	المُضلل الثانوي	نادر الاستخدام
	المُضلل الثالث	نادر الاستخدام
القصد/النية	الغرض المستخدم	نص مجاني (جملة أو جملتان قصيرتان)
	الفئة	على سبيل المثال: مدنية أو اجتماعية أو اقتصادية أو عسكرية

الجدول 1. نسخة مُبسّطة من مخطط تصنيف حملة التّضليل السياسية من تصميم بروكينغ وآخرون (2020).

في هذا العمل يعرّف بروكينغ وآخرون التّضليل السياسي بأنه: «تضليل له غرض سياسي أو لصيق بالسياسة» مما يشمل «التّضليل في سياق الانتخابات أو الاحتجاجات أو العمليات العسكرية إضافةً إلى «الظاهرة واسعة الانتشار ألا وهي: العناوين السياسية المثيرة للانتباه ‹clickbait› التي تنتشر لتحقيق مكاسب مالية شخصية».

ويعرّف إطارهم حملة التّضليل السياسي بأنها: «تضليل سياسيٍ يدل على وجود إعداد مُسبق وغرض مُبطّن». وهم يؤكّدون، أولاً، أنّ الأهداف قد لا تكون دائماً واضحةً، على الرغم من أنها موجودة لا محالة؛ وثانياً، أنّ الحملات ذات الأهداف المتغيرة قد تصبح مختلفةً عن بعضها البعض (بما معناه أنّ الحملة تصبح حملةً جديدةً متى تغيرت الأهداف). وعلاوةً على ذلك، يذكرون أنّ حملات التّضليل السياسي تنطوي دائماً تقريباً على ما يسمونه «تضخيم المُحتوى». هذا المفهوم، الذي سيُناقش بمزيد من التفصيل في القسم التالي، يطلق عليه «كيلر وآخرون» (2019) اسم «الدعاية الزائفة السياسية» أو «political astroturfing» ، فيما يطلق عليه فيس بوك اسم «السلوك الزائف/المُفتعل المُعد له» «coordinated inauthentic behavior» كما يرد في (Gleicher, 2018a)، ويُشار إليه باعتباره سمةً من سمات الدعاية الشعبية الزائفة في استهدافه للصحافيين بمعلومات مضللة مصممة «لتضليلهم أو إعطائهم معلومات مغلوطة أو إرباكهم أو تعريضهم للخطر» كما يؤكد «بوزيتي» (Posetti, 2013) ولا تُشكّل كل هذه الحالات تضليلاً في حد ذاتها، ولكن هناك تداخل واضح لأن الهدف هو خلق «وهم بأن المعلومات تتمتع بالتوافق أو الشعبية»، وفي بعض الحالات، إلحاق الأذى. وقد حاول بعض الباحثين تمثيل هذا التفاعل المعقد من خلال «مصفوفة أضرار التّضليل»، التي تشمل الاستقطاب والتطرف من جهة والدعاية والإعلان من جهةٍ أخرى (Frau-Meigs, in press).

ويتيح هذا الإطار أساساً لمقارنة مختلف أنواع المعلومات المُضلّلة كما ويسمح بتمثيل المعلومات المرجعية المفصلة. ومن شأن فهم السياق الظرفي، على غِرار وجود نزاع عسكري، أو مستويات الاستقرار السياسي أن يساعد في التقييم قصير وطويل المدى، وعلى توفير الحلول المناسبة. ومع ذلك، فإنه قد يعني أن بعض العوامل قد تكون غير معروفة أو من دون صلة بالموضوع. وكما في حال

الأُطُر الأخرى التي تمت مناقشتها، فإن مفهومي القصد والإسناد ليسا دائماً واضحين. وعلى قدر أهمية التّضليل خلال الحملات السياسية، فإن هذه الدراسة تضع في اعتبارها الصورة الأوسع التي تتضمن إمكانية وجود أكاذيب غير مقصودة (كما في المسائل الصحية)، وبالتالي تحافظ على تركيزها على استجابات أوسع من تلك التي تتطرّق للقضايا السياسية فقط.

3.1.3 تأثير المعلومات

على غِرار توصيف حملة التّضليل السياسي، ينظر «دليل العاملين في مجال التواصل» (Pamment et al., 2018) أو the Handbook for Communicators إلى المعلومات المُضلّلة في سياقها الأوسع كجزء من «أنشطة التأثير» ومن منظور وضع السياسات (في حالة هذا الدليل، الحكومة السويدية). ويفكك الإطار المطروح في الدليل أنشطة التأثير التي تستخدمها القوى الأجنبية، مع التركيز على الاستراتيجيات الخطابية، والتقنيات ومُخطّطات التأثير، لتمكين صُنّاع السياسات من تحديد وفهم ومواجهة هذه الأنشطة والحملات التي تزداد تطوراً يوماً بعد آخر. وتُركّز هذه المقاربة بشكلٍ خاص على حماية «الحوار الديمقراطي» في المُجتمع، الذي يعرّف بأنه «الحق في النقاش المفتوح، والحق في تكوين رأي معيّن بحُريّة، والحق في حُريّة التّعبير». وفي ضوء ذلك، فإنّ الدليل يعتبر أن أساليب تعزيز الصُّمود الاجتماعي، على غِرار إعلام الجمهور وتثقيفه، هي أساس مُكافحة المعلومات المُضلّلة وأنشطة التأثير، ويُركّز على الجهات المعنية بالتواصل مع الجمهور داخل الحكومات والمؤسسات الحكومية.

ويرتبط «تأثير المعلومات» في هذا الإطار ارتباطاً وثيقاً بالمعلومات المُضلّلة ويعرّفه «بامنت وآخرون» بأنه «تقنية تستند إلى توزيع معلومات خاطئة بهدف التّضليل والخداع». ويؤكّد المؤلفون أنّ من يقوم «بأنشطة التأثير» هو قريب جداً من الحملات الإعلانية (المشروعة تماماً) التي تحاول التأثير على الناس لشراء المنتجات، على سبيل المثال. وهم يجادلون بأن مفهوم الانفتاح هو بالضبط ما يميّز بين المسألتين: فالإعلان والعلاقات العامة ينبغي أن تكون شفافةً في دوافعها، وأن تتبع قواعد واضحة. من ناحية أخرى، ينطوي تأثير المعلومات على نشر المُحتوى الكاذب بشكل سري ومخادع. في هذا الصدد، يتقاطع نهج «بامنت» إلى حد كبير مع الاستخدامات الأوسع نطاقاً لمصطلح «حملات المعلومات» الذي يشير إلى مزيج من السُّلوكيات المُفتعلة المُعد لها (على غِرار الحسابات الزائفة والسلوكيات الخفية) وهي ظاهرة أوسع من عمليات النشر العسكرية أو الجيوسياسية.

> **”نظراً إلى أن المجتمعات مبنية على الثقة، فإن «تأثير المعلومات» المُخادع يُقوّض المبدأ الديمقراطي المتمثل في حق الجمهور في المعرفة والنفاذ والوصول السّلس إلى المعلومات“**

ونظراً إلى أن المجتمعات مبنية على الثقة، فإن «تأثير المعلومات المُخادع» يُقوّض المبدأ الديمقراطي المتمثل في حق الجمهور في المعرفة والنفاذ والوصول السّلس إلى المعلومات، بل قد يُزعزع استقرار الديمقراطية عن طريق تعكير المياه الإعلامية لدرجة قد يصبح من المستحيل التمييز بين المعلومات الدقيقة والأكاذيب وبين الصحافة ذات المصداقية من جهة، والدعاية من جهة أخرى، مما يقوّض الثقة في المعلومات التي تهدف إلى تحقيق المصلحة العامة بشكل كبير. وفي هذا الصدد، فإنّ مفهوم «تأثير المعلومات» يتقاطع أيضاً في جزء منه مع مفهوم أزمة المعلومات المُضلّلة الذي وضعته منظّمة الصحة

العالمية[19] والذي يُشير إلى «سيل معلومات– منها الصحيح ومنها غير الصحيح – في ظل جائحة؛ ممّا يصعّب على الأفراد العثور على مصادر جديرة بالثقة وإرشادات موثوقة عند الحاجة»[20].

وتتضمّن نظرية تأثير المعلومات التي اعتمدها «بامنت وآخرون» ثلاثة أجزاء: الإدراك، وتحديد المعلومات، والتصدي لها.

يتضمن الإدراك فهم تكوين حملات المعلومات وعملية تكوين الرأي. وفي ضوء ذلك، يتميّز تأثير المعلومات بثلاث صفات رئيسية: الخداع والقصد وإحداث الاضطراب. لكن لابد من الإشارة إلى أنه ليس من السهل أو من الممكن دائماً تحديد وجود هذه الجوانب؛ مما يدل على وجود فجوة كبيرة في هذه النظرية. وكما سبق وذُكر؛ قد يكون من الصعب تحديد القصد، أو على الأقل إسناده، كما ويصعب قياس مدى وأثر الاضطراب الناتج.

وتستند عملية تحديد «تأثير المعلومات» على فكرة السرديات الاستراتيجية، التي يمكن اعتبارها تلاعُباً مُتعمّداً ببعض المعتقدات الراسخة على غِرار كُرويّة الأرض[21]. وتختلف الفئات المستهدفة هنا عن تلك التي تم تعريفها في الأُطُر الأخرى كإطار «ديرخشان» (2019) و«بروكينغ وآخرين» (2020). فالفئات المستهدفة هنا هي دائماً الجمهور أي الجمهور العام، ومجموعات اجتماعية عِرقيّة (على سبيل المثال: مجموعة دينية)، ومجموعات سيكوغرافية (أي تلك التي لها سمات نفسية وشخصية محددة).

وفي إطار «بامنت»، يتم تعريف التّضليل بشكل أضيق بكثير من التعريف المُعتمد في هذا التقرير. فيصنّف على أنه تقنية مختلفة عن التقنيات التي تنطوي على استغلال تقني يشمل روبوتات الإنترنت، والتزييف الفج، وشبكات الدُّمى. وتُعتبر هذه التقنيات بدورها مختلفةً عن فئة «الهُويات المُخادعة» التي تشمل ما تسميه «المُحتوى الإعلامي الزائف»[22] و«المُحتوى المُزوَّر». أما الفئات الثلاث الأخرى أي القرصنة الاجتماعية والمعرفية، والخطاب الخبيث، والإجراءات الرمزية، فترتبط بشكل أقل تشدداً بالتّضليل، وتشمل مفاهيم التحيُّز على غِرار غُرف الصدى filter bubbles، وتكتيكات التخويف والخروج عن الموضوع الأصلي strawman tactics، والتسريب leaking والقرصنة hacking، على التوالي. من ناحية أخرى، يتم تصنيف التقليد السّاخر والتهكُّم (بشكل إشكاليّ) على أنهما من المعلومات المُضلّلة. وتقر الدراسة على عكس ذلك بأنّ التّضليل غالباً ما يكون مُدبراً، لكنه ليس تقنيةً في ذاته بل إنه يستخدم تقنيات على غِرار التكنولوجيا والهُويات المُخادعة. وهذا ينطبق أيضاً على تحليل «فرانسوا» Francois

---

19 https://www.merriam-webster.com/words-at-play/words-were-watching-infodemic-meaning
https://www.who.int/teams/risk-communication/infodemic-management

20 https://www.who.int/news-room/events/detail/2020/06/30/default-calendar/1st-whoinfodemiology-conference

21 تجدر الإشارة إلى أن المعلومات المُضلّلة، كما تفسر في هذه الدراسة، يمكن أن تعتبر نوعاً من أنواع تأثير المعلومات. لكن ذلك لا يعني بالضرورة أن جميع السرديات الاستراتيجية تُساوي بين المعلومات المُضللة، ولا أن جميع السرديات الاستراتيجية مخادعة بشكل أساسي.

22 مُلاحظة من المحررين: يطرح مصطلحا «الأخبار الزائفة» و«المُحتوى الإعلامي الزائف» إشكالية، ويجب تجنبهما حيثما أمكن ذلك لأنهما يُستخدمان غالباً كسلاح في «ترسانة» التّضليل وذلك في محاولة لتشويه سمعة الصحفيين والتقارير الإخبارية التي تضعها الجهات الفاعلة التي تسعى إلى نشر مفهوم صحافة المُساءلة. لمزيد من المناقشة، الرجاء مُراجعة دليل اليونسكو حول: «الصحافة والأخبار الزائفة والتّضليل» (قابل للتحميل مجاناً بلغات متعددة على الرابط التالي: (https://en.unesco.org/fightfakenews

(2019)، الذي يعتبر أنّ السلوكيات، بما في ذلك السلوكيات المُفتعلة (والجهات الفاعلة الزائفة)، من الصفات الأساسية التي تميّز ما ينبغي اعتباره معلومات مُضلّلة. ولكن في حين تُقر هذه الدراسة بأنّ هذه الصفات مهمة؛ تؤكّد على أن العديد من حالات التّضليل لا تتضمّن هذه الصفات.

وكما ذكر عديدون، يُواجه «إطار بامنت» التحدي المتمثل في قدرته العملية على التمييز بناءً على افتراضات حول الدافع والقصد. وينطبق هذا التحدي أيضاً على تفسير السلوكيات على أنها مقياس للدوافع، حيث أن هناك اختلافات معقدة بين الشخص الذي يشارك مُحتوى غير صحيح ظناً منه أنه صحيح ومفيد، ووكيل يضخم هذا المُحتوى، ومُحرّض يعمل في إطار استراتيجية أوسع. من ناحيةٍ أخرى، فإن تقييم «بامنت وآخرين» يتجنّب المشكلة المحتملة التي يطرحها مفهوم «المعلومات الضارة»، حيث يقر بأنه لا يمكن اعتبار كافة الروايات المقنعة أو الاستراتيجية من المعلومات المُضلّلة.

وأخيراً، من حيث استراتيجيات مُكافحة تأثير المعلومات، يقترح «بامنت وآخرون» أربع فئات، وفقاً لترتيب زمنيّ. فالاستجابتان الأوليان هما: التقييم والإعلام، وهما تقنيتان قائمتان على الوقائع. وتتبعهما استجابتان قائمتان على المناصرة وأخيراً استجابة قائمة على الدفاع. وقد تنطوي الخُطوة الأولى، وهي تقييم الوضع، على طرق فعالة كالتحقُّق من الوقائع والتحقيق في شفافية المعلومات. ويتضمن الإعلام خُطوات على غِرار الإدلاء بتصريحات للإشارة إلى وجود مشكلات معينة وتصحيح الأخطاء المتعلقة بالوقائع. وتوصف المُناصرة بأنها استخدام آليات معينة مثل الحوار وتيسيره. أما الدفاع فهو المرحلة الأخيرة التي تنطوي على حجب المعلومات المُضلّلة والإبلاغ عنها وإزالتها. وعلى الرغم من أن «إطار بامنت» لا يقترب من النطاق الواسع للاستجابات التي تُغطيها هذه الدراسة، فهو يتميّز بتسليط الضوء على الروابط بين الإدراك وتحديد المعلومات والاستجابة.

2.3 البحوث التجريبية والتطبيقية

ينتقل هذا القسم من الأُطُر النظرية التي تحاول تحديد وتصنيف الأنواع المختلفة من المعلومات المُضلّلة، وفي بعض الحالات، الاستجابات المحتملة لها، ليُركّز على البحوث التجريبية والتطبيقية وينظر في بعض الاتجاهات الرئيسية؛ وبأمثلة عن دراسات حالة محددة.

ويركز تقرير «المسح العالمي للتلاعب المنظم بوسائل التواصل الاجتماعي» الذي أعده «برادشو وهاورد» (2019) على التلاعب بوسائل التواصل الاجتماعي على يد الحكومات والأحزاب السياسية. ويحلل هذا التقرير اتجاهات ما يسمى بـ«الدعاية الحاسوبية»، ويبحث في الأدوات والقدرات والاستراتيجيات والموارد. ويُظهر المسح أنه في السنوات الأخيرة، انتشرت الأدلة على تنظيم حملات للتلاعب بوسائل التواصل الاجتماعي في جميع أنحاء العالم، مع ارتفاع عدد البلدان المُشاركة في هذا التلاعب بنسبة 150% في غضون عامين. ففي عام 2019، ظهرت أدلة على وجود هذه الحملات في 70 دولة، مُقارنةً بعام 2018 الذي بلغ فيه هذا العدد 48 دولة، وعام 2017 الذي بلغ فيه هذا العدد 28 دولة فقط. كما وتبيّن أن فيس بوك هو وسيلة التواصل الاجتماعي الأكثر استخداماً في هذه الحملات.

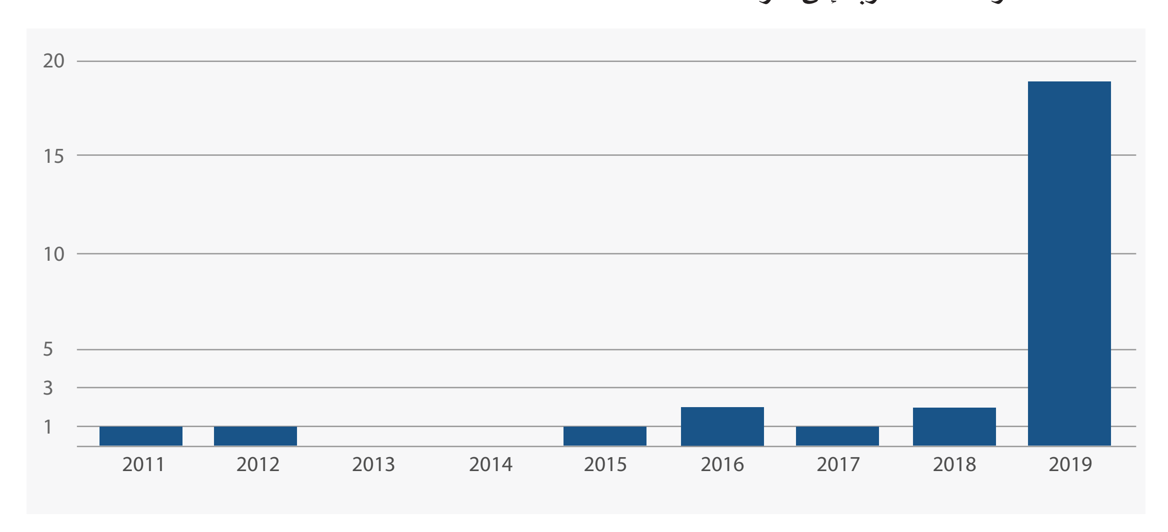
ويطرح مارتن وشابيرو (Martin & Shapiro, 2019) أيضاً نظام تصنيف مُفصَّل لمبادرات «التأثير الأجنبي» عبر الإنترنت يقارن بين الخصائص والتكتيكات والخطاب والمنصات التي تستخدمها مختلف الجهات التي تشن هذه الهجمات. وقد حاولت بعض الدراسات التعمق في الدوافع الكامنة وراء هذه الأنواع من المبادرات، ولكن هذه الدراسات تقتصر على دراسات حالة فردية خاصة بكل بلد. أما «أونغ وكابانيس Ong and Cabañes, 2018» فيبحثان من منظور «مُستتير عِرقياً» في دوافع وسلوكيات من يتم تسخيرهم لإنتاج معلومات مضللة مُعد لها وللتلاعب بوسائل التواصل الاجتماعي في الفلبين، بينما يقوم «تشاتورفيدي Chaturvedi, 2016» بالتحقيق في قضايا مماثلة في الهند.

ومع ذلك، وعلى الرغم من هذه التقارير وغيرها من التقارير التي تناقش هذه الأنماط من التّضليل السياسي المنظم و«حملات التأثير»؛ لا يزال هناك نقص في البحوث المتعمقة المُعد لها حول هذه الظاهرة، خاصةً على مستوى أعلى من مجرد دراسة حالة أو بلد محدد. ويمكن لهذه الحملات أن تؤثر على الناس لدرجة تدفعهم إلى تغيير تصويتهم أو شراء منتجات معيّنة أو تغيير تصوراتهم، مما له في بعض الأحيان تداعيات وخيمة على الديمقراطية أو الصحة العامة. وقد تم تعريف مفهوم «العلاقات العامة الهدّامة/المظلمة» بأنه «تلاعب يتمّ على نطاق واسع بهدف الربح ومن دون الاكتراث بالأضرار التي قد يلحقها بكوكب الأرض أو بدولةٍ ما أو بسلامة الأفراد»[23] مما يؤدي إلى نشوء صناعة عالمية مؤلفة من شركات علاقات عامة وتسويق؛ تقوم بشراء خدمات تستخدم حسابات زائفة وسرديات كاذبة لنشر التّضليل عبر أنظمة تلاعب عبر الإنترنت من طرف إلى طرف (Silverman et al., 2020).

[23] تعريف «روب إندرلي»، مدير في «مجموعة إندرلي»، ورد كاقتباس في مقالة لـ«ريتشارد أديكاري» Richard Adhikari صدرت في مجلة E-commerce Times وعنوانها «شركات العلاقات العامة المظلمة تملأ جيوبها عن طريق نشر المعلومات المُضلّلة» بالإنكليزية Black PR’ Firms Line Their Pockets by Spreading Misinformation by Richard Adhikari يمكن مُراجعتها على الرابط التالي: https://www. ecommercetimes.com/story/86444.html

وقد سعت العديد من الدول في جميع أنحاء العالم إلى تجريم إنتاج وتوزيع المعلومات المُضلّلة من هذا النوع (Adhikari, 2020)، على الرغم من اختلاف تعريف ما هو مقبول في هذا المجال اختلافاً كبيراً. من الناحية العملية، قد لا يكون من السهل على سُلطات إنفاذ القانون أو الشركات نفسها العثور على مصادر التّضليل وإثبات نيتها. ويُشير «أديكاري» إلى أنّ فيس بوك قد حاول الحد من ممارسات نشر المعلومات المُضلّلة هذه، فحظر في عام 2019 عدداً من شركات العلاقات العامة المظلمة لمحاولتها التأثير على الانتخابات، أو لما يسميه «السلوك المُفتعل المُعد له» في مختلف البلدان. ومع ذلك، لا تزال هذه الأنواع من الأنشطة واسعة الانتشار، ومن السهل إنشاء شركات جديدة تروّج لمثل هذه الخدمات.

حملات المعلومات المنسوبة إلى شركات العلاقات العامة:



الشكل 3. مصدر الرسم البياني Buzzfeed (Silverman et al., 2020)

لقد أدى تحليل أجرته محطة BBC لقصص مضللة روّجت خلال جائحة فيروس كورونا إلى تصنيف الجهات الفاعلة المشاركة في اختلاق معلومات مضللة في سبع فئات (Spring, 2020). ومن التحقيقات الصحفية الأخرى الجديرة بالذكر والتي تعطي نظرةً ثاقبةً حول الوكلاء والمُحرّضين تحقيقات أجراها كل من Buzzfeed وCodatory وRappler على سبيل المثال (Dorroh, 2020) كما هو وارد لاحقاً بمزيد من التفصيل في القسم 7.1. غير أنه عند كتابة هذا التقرير؛ كانت هناك نُدرة في الدراسات الأكاديميّة المُفصَّلة حول هذه الظاهرة، ولم يكُن من الواضح ما هي الطرق التي تسمح بالوقاية منها من المصدر.

في حين أنّ «حملات التأثير» ليست جديدة في حد ذاتها، فإنّ انتشار هذه الحملات بشكل واسع في عام 2019، كما هو موضح أعلاه، يستلزم اهتماماً عاجلاً. من ناحية أخرى، تجدر الإشارة إلى أنه لا يمكن أن تُشمل كافة «حملات التأثير» بالتوصيف المستخدم في هذه الدراسة للمعلومات المُضلّلة، بمعنى أنّ بعض هذه المبادرات قد لا يستخدم مُحتوى كاذباً أو مضللاً، وقد لا يقوم على سلوك مُفتعل. ومن أهم الحالات البارزة الحديثة حالات ترتبط بآليات على غِرار «إثارة النّعرات العنصرية» (Neate, 2017) و«السُّلوك المُفتعل المُعد له» (Gleicher, 2019a) وتفتح الباب أمام عدد من الاحتمالات لفائدة استخدامها للمحتويات المُضلّلة. ويمكن تنظيم بعض الحملات المُعد لها باستخدام محتويات صحيحة، وسلوكيات شفافة، وجهات فاعلة موثوقة، كما في أنشطة المناصرة التي تنظمها بعض مجموعات المجتمع المدني، وحملات التواصل التي تجريها الحكومات في مجال الصحة العامة، ومبادرات العلاقات العامة التي تقوم بها الشركات. وبالتالي،

لابد من تناول موضوع حملات التأثير المُعد لها عبر مقاربة تراعي الفوارق البسيطة عند البحث في الأماكن والطرق التي تتقاطع فيها حملات التأثير مع المُحتوى الكاذب والمضلل والضار.

1.2.3 الأُسُس الاجتماعية والنفسية

تسعى مجموعة من البحوث حول المعلومات المُضلّلة إلى وضعها في سياقها الاجتماعي والنفسي للسماح بتحديد وفهم الاستجابات المناسبة. وحتى لو كانت بعض آليات التّضليل جديدةً، فبإمكان الاستجابات أن تسترشد بالبحوث التي أجريت على مدى عقود حول الإدراك البشري. وكما سيتم مناقشته في مكان آخر في هذه الدراسة، قد يكون من الصعب إقناع الأشخاص الذين يرغبون في تصديق معلومة معيّنة بأن هذا المُحتوى خاطئ بالفعل أو أنّ «حقيقةً» خاطئةً قد تُحدث فرقاً في فهمهم للمشهد الأوسع. وحتى لو تمكّن المحققون من الوقائع من دحض المعلومات الكاذبة، فقد أظهرت الأبحاث أنه قد يكون من الصعب جداً تغيير قناعات الأفراد حول المفاهيم الخاطئة، خاصةً إذا كانوا يعتقدون أن هناك نواة حقيقة داخل الزيف. (Johnson & Seifert, Research Context and Gaps 51 1994; Nyhan & Reifler, 2010)

وكما كتب الخبير الاقتصادي «جي. كاي. غالبريث» J.K. Galbraith ذات مرة: «في كل مرة يُخيّر فيها الناس بين تغيير رأيهم وإثبات عدم الحاجة إلى ذلك، ينشغل جميعهم تقريباً بتقديم إثبات لعدم حاجتهم إلى تغيير رأيهم» (Galbraith, 1971) فالتكرار والخطاب أداتان قويتان: والناس يميلون لتصديق المعلومات عندما تتكرر عليهم مراراً وتكراراً (Zacharia, 2019). والأهم من ذلك، وفقاً لـ«أفرون وراج, Effron & Raj 2019» أنّ هذا التعرض المتكرر للمعلومات نفسها يقلل من الحواجز الأخلاقية التي قد تثني هؤلاء الأشخاص عن مشاركة المُحتوى، سواءً أكانوا يصدقونه أم لا.

وقد أظهرت الأبحاث طويلة الأمد في العلوم السياسية قوة الخطابة مراراً وتكراراً (Kroes, 2012; Grose and Husser, 2008): فمن الأرجح أن تؤثر الخطب المنمقة لُغوياً التي يلقيها المرشحون للانتخابات على الناس فتدفعهم إلى التصويت لصالحهم. وينطوي هذا التنميق اللُّغوي على عرض رسالة ما –بغض النظر عن محتواها– بطريقة بلاغية مصممة خصيصاً لتحدث صدىً عاطفياً. غير أن إحدى الخلاصات تؤكد على أن التنميق اللُّغوي (أي تقديم الرسالة بطريقة بلاغية معينة، لا تغيير الرسالة نفسها) من المرجح أن يقنع الحاصلين على تعليم عالٍ بأمر ما، ولكنه لا يغيّر رأي من لا يتمتع بالتعليم العالي (Grose and Husser, 2008). وفي حين لا ينبغي المساواة بين الخطابات الانتخابية في حد ذاتها وبين التّضليل، فإنّ هذه الخلاصات تشير إلى أنّ للمعلومات المُضلّلة أثر أقوى عندما تقترن بأبعاد غير إعلامية (أي بعوامل عاطفية) مقارنةً بأثرها عندما تُعرض وحدها. ويذهب البعض أبعد من ذلك، ليضعوا الجوانب المتعلقة بالتّضليل في إطار «الحرب النفسية» (Szunyogh, 1955, cited in Cordey, 2019). ويقدّم «غوادانيو وغوتيري» (2019) لمحةً عامةً عن البحوث في علم النفس والعلوم السياسية التي تتمحور حول انتشار المعلومات المُضلّلة. ويستعرضان عدداً من العوامل الاجتماعية والسياقية والفردية التي تساهم في هذا الانتشار. وبالتركيز على انتشار وتأثير «الدعاية المظلمة» عبر الإنترنت بشكل خاص، قاما بدراسة العناصر الاجتماعية كالتفاعلات عبر الإنترنت والتطوُّرات التكنولوجية التي تؤثر على ذلك. وهما يضعان المعلومات المُضلّلة في سياق عوامل أخرى تتعلق بوسائل الإعلام قد تسهم أو تدفع نحو انتشار وتأثير المعلومات المُضلّلة. غير أن بحثهما يركز فقط على دراستين فرديتين محددتين (الولايات المتحدة[24] وإستونيا). وفي حين يسجلان اختلافات

---

[24] سحبت الولايات المتحدة عضويتها من اليونسكو في تشرين الأول، أكتوبر 2017.

بين هاتين الحالتين فلا يمكن بالضرورة تطبيق نتائج هذا البحث على موقع جغرافي أو ظرف أوسع. إلى جانب هذه المفاهيم، قد يكون من المفيد فهم بعض الأسباب التي تدفع بالأفراد إلى تصديق المحتويات الكاذبة، وإلى مشاركتها مع آخرين على الرغم من شكهم أو علمهم بعدم صحتها. وفي هذا الإطار، أجريت عدة دراسات حول «سيكولوجيا المعتقد»، تؤكد أنّ العلوم السلوكية يجب أن تلعب دوراً رئيسياً في إرشاد الاستجابات الساعية إلى التصدي للمعلومات المُضلّلة (Lorenz-Spreen et al., 2020). ويقوم «ليفاندوفسكي» Lewandowski بدراسة نظريات المؤامرة بشكل خاص على غرار تلك التي تتمحور حول جائحة فيروس كورونا. فيعتبر أنه خلال الأزمات؛ يمر الاشخاص عادةً بمراحل مختلفة من الإنكار تتضمن عدم الاعتقاد بوجود أزمة من الأساس، ثم إلقاء اللوم على الآخرين، وعدم الاعتقاد بنجاعة أيّة حلول. وكل هذه المراحل تؤدي عادةً إلى دعم نظريات المؤامرة هذه (Cook et al., 2020).

وفي البلدان التي تسيطر فيها السُّلطات الحكومية على وسائل الإعلام التقليدية إلى حد كبير أو بشكل كامل، غالباً ما ينتشر انعدام الثقة في هذه الوسائل، لا سيما عندما يكون ذلك مرتبطاً بمشكلات قديمة أو حالية مثل الفصل العنصري والفساد. في مثل هذه البلدان، غالباً ما يعتبر الجمهور أنّ «راديو ترو توار» (حرفياً، راديو الرصيف) (Ellis, 1989) وأشكال أخرى من وسائل الإعلام السرية هي أكثر موثوقيةً من المصادر الرسمية للمعلومات (Wasserman, 2020). وقد بيّنت دراسة لـ«واسرمان» أُجريت في أفريقيا جنوب الصّحراء الكبرى (Wasserman & Madrid-Morales, 2018) تدني الثقة في وسائل الإعلام بين المواطنين وارتفاع التعرض للتضليل. كما تبيّن أن المواطنين في هذه الدول غالباً ما يُساهمون في انتشار المعلومات المُضلّلة حتى مع علمهم بعدم صحتها بدرجة أكبر بكثير من المواطنين في الولايات المتحدة الأميركية. ويبرز هذا الاستنتاج أنه لابد من أن نعرف إلى أي مدى على استراتيجيات مُكافحة التّضليل أن تتعدى استراتيجيات التثقيف والدراية الإعلامية البسيطة لتتمكن من التصدي للأسباب الجذرية التي تقف وراء انعدام الثقة.

ويُشير عمل «واسرمان وإليس»، من بين أمور أخرى، إلى أنّ أسباب مشاركة معلومات كاذبة عن سابق علم بزيفها قد تكون مرتبطةً بمفهوم الولاء الجماعي. بعبارة أُخرى، يبدو أن مفهوم الحقيقة أقل أهميةً من مفهوم التضامن. فما دامت المعلومات تتماشى مع نظرتنا إلى العالم، غالباً لا نقوم بالتحقُّق من صحتها. من ناحية أخرى وُجدت دراسة أجراها «بانيكوك وراند 2019» Pennycook & Rand أن هناك فرق كبير بين قدرة الأفراد على التمييز بين المعلومات الصحيحة والمعلومات الخاطئة من جهة واحتمال مشاركتهم تلك المعلومات من جهة أخرى. هذا يعني بعبارة أخرى أنّ الأفراد لا يشاركون فقط المعلومات التي يعتقدون أنها صحيحة. ويتضح من كل هذه النتائج أن «البخل المعرفي»[25] والتحيز المعرفي يلعبان دوراً مهماً في تصديقنا للمعلومات الخاطئة ومشاركتنا إياها خاصةً في بيئة غنية بالمعلومات. كذلك، يتضح أننا نتأثر بالاستدلال على غرار التأييد الاجتماعي. لذلك لابدّ أن تؤخذ هذه العناصر في الحُسبان عند تقييم الاستجابات للمعلومات المُضلّلة.

من أجل التصدي للتضليل بفعالية، من المهم أيضاً أن نفهم بعض الأسباب التي تدفع بالأشخاص إلى عدم تغيير آرائهم حتى عندما يُواجَهون بأدلة على عدم صوابها. يناقش «هانز روزلينغ» Hans Rosling

[25] إنّ «البخل المعرفي» مفهوم يعتبر أننا نفضل اتخاذ قرارات سهلة تتماشى مع تصوراتنا المسبقة، وأننا وقد ننسى تفاصيل معيّنة (أنّ المعلومات قد تم كشفها سابقاً على سبيل المثال) https://firstdraftnews.org/latest/the-psychology-of-misinformation-why-were-vulnerable/

الفكرة القائلة بأن للأفراد عادةً عدداً من المفاهيم الخاطئة السلبية حول العالم (مثل: متوسط العمر المتوقع في البلدان الفقيرة، أو معدل الوفيات الناجمة عن الكوارث الطبيعية). فحتى عندما يواجَه هؤلاء بأرقام تدحض هذه المفاهيم، لا يتقبلونها بسهولة (Rosling, 2018). وهو يعزو ذلك إلى ثلاثة عوامل: الحنين الزائف (ذكريات خاطئة عن الماضي تُصوّره على أنه أفضل مما كان عليه في الواقع)؛ التقارير الانتقائية من قبل الصحفيين (التي تُسلّط الضوء مثلاً على القصص السلبية تماشياً مع القيم الإخبارية التقليدية التي تعطي الأولوية لتصوير المعاناة والفساد والخطأ)، والشعور بأنه من غير المناسب إلى حد ما الحديث عن أي تحسن طفيف خلال الأزمات. وكثيراً ما يستغل انتشار المعلومات المُضلّلة هذه المعتقدات ويتلاعب بها، لا سيما في ما يتعلق بالأزمات والصراعات والكوارث الطبيعية. وفي حين يُشجّع «روزلينغ» مفهوم التثقيف العام للتصدي للمعلومات المُضلّلة؛ ليس هناك من أبحاث تقيس فعالية هذه الاستراتيجية، خاصةً بالنظر للنتائج التي توصل إليها «روزلينغ».

ويركز تقرير منظمة Witness المعنون «Ticks or It Didn't happen» على الاستجابات للمعلومات المُضلّلة من منظور أخلاقي في المقام الأول (Witness Media Lab, 2019). ويدرس التقرير إحدى التقنيات الأساسية المستخدمة لتعقب أصالة الصُّور (مقاربة «نقطة الالتقاط» point-of-capture على مستوى الكاميرا)، فيستعرض 14 معضلةً متعلقةً بهذه التقنية، حيث أنّ الهيكلية التي تبيّن الأصالة تعتبر من الاستجابات للمعلومات المغلوطة وللمعلومات المُضلّلة وللتلاعب بالمواد الإعلامية. وتشمل هذه المعضلات معضلات تقنيةً تتمحور حول مسائل النفاذ، والخصوصية، والمراقبة، والخيار الحكومي المشترك، ومخاوف تتعلق بوضع درجات على المصداقية قد تكون تبسيطية بشكل مفرط أو صعبة التقييم. ويركز التقرير على استخدام إطار معضلات كولينجريدج (Collingridge, 1980) حول القدرة على التأثير في النظم التكنولوجية والقيام بذلك في وقت مبكر بما يكفي لضمان أن تعبر هذه النظم عن قيم حقوق الإنسان وإلا سيتم تجاهل هذه الحقوق متى تطورت هذه النظم وأصبحت واسعة النطاق. وينطبق هذا المنظور أيضاً على مجموعة من المقاربات التكنولوجية للمعلومات المُضلّلة والتي قد تعطي أو لا تعطي الأولوية لحُريّة التّعبير أو غيرها من قضايا حقوق الإنسان.

2.2.3 التركيز على النواقل

إلى جانب المفاهيم المتعلقة بالإقناع ومُكافحة المعتقدات الخاطئة، على الاستجابات للمعلومات المُضلّلة أن تأخذ في الاعتبار طبيعة المعلومات المُضلّلة، ومن تستهدف (كما نوقش أعلاه)، ودور آليات النقل، أو ما يُعرف بالنواقل. وتلعب هذه النواقل دور الوسيط بين إنتاج واستخدام المعلومات المُضلّلة مما يتيح تداولها بطرق مختلفة وعلى مختلف المستويات. ولابد من فهم الأنماط التي تدخل في هذا القسم من دورة المعلومات المُضلّلة لإرشاد الاستجابات ليس فقط في ما يخص نقل المعلومات بل أيضاً في ما يتعلق بالاستراتيجيات التي تستهدف الإنتاج الأساسي والاستخدام اللاحق للمعلومات المُضلّلة.

وهناك ثلاث آليات رئيسية تسمح بنقل المحتويات الخاطئة. أولاً، قد يهدف التّضليل إلى تعطيل عمل وسائل الإعلام الإخبارية أو استغلالها للوصول بشكل غير مباشر إلى أهدافه، سواء كانت من الجهات الحكومية أو غير الحكومية. وتشكل وسائل الإعلام المستولى عليها والصحفيون المشكك في نزاهتهم أو ضعف قدرات التحقُّق أوجه ضعف يتم استغلالها. من ناحية أخرى، قد يكون التّضليل استغلالاً مُعدّاً له (وغالباً بصورة تلقائية، ولكن ليس دائماً) و/أو لعباً على منصة إلكترونية بهدف الوصول إلى الجمهور

(أي استهداف جزئي لطبيعة نموذج الأعمال في هذه المنصة ونطاق وصولها). وفي حالات أخرى، تستهدف المعلومات المُضلّلة الجمهور العريض في المقام الأول بغرض تحقيق نشرها أكثر، معتمدةً على قدرة الجمهور على تحقيق انتشار المعلومات على نطاق واسع باستخدام أطراف ثالثة تلعب دور الوسيط من نظير إلى نظير للوصول إلى جمهور أكبر. في كل من هذه الحالات على الاستجابات استهداف الآلية ذات الصلة في المقام الأول (وسائل الإعلام، شركات الإنترنت، والجمهور).

3.2.3 الدفاع عن القيم العامة في «مجتمع المنصة»

في حين قد تلعب وسائل الإعلام الإخبارية والجمهور دور نواقل المعلومات المُضلّلة، ينظر هذا القسم بمزيد من التفصيل في دور شركات التواصل عبر الانترنت (والتي غالباً ما يشار إليها باسم «منصات») كقنوات وأدوات تُستخدم لتضخيم ونشر المعلومات المُضلّلة. وقد أدى صعود التكنولوجيات الرقمية إلى تزايد أهمية البيانات فأصبحت هذه الشركات معاقل جديدة للسيطرة والربح حيث أنه بإمكانها تخزين كميات هائلة من المُحتوى والتلاعب بها مما قد يعني قدرتها المحتملة على التلاعب بالجماهير. وقد أدى هذا بدوره إلى صعود لاعبين مهيمنين (Srnicek, 2017) وله تداعيات مهمة على إنتاج ونشر واستخدام المعلومات وموثوقيتها. وتواكب مبادرة المنظمة غير الحكومية Public Knowledge التي تعمل على الموقع /https://misinfotrackingreport.com، سياسات وممارسات عدد من الشركات التي تتعامل مع هذه التحديات. كما وتقوم حركة المجتمع المدني آفاز Avaaz بتعقب مظاهر السرديات المُضلّلة لمواضيع محددة، وتقييم أداء الشركات في مُكافحة هذا المُحتوى[26].

ويمكن القول إلى حد ما، أنّ ما يضعف المنصات الرقمية ويجعلها قنوات للتضليل هو نموذج عملها، ولكنها قد تكون أيضاً عوامل تمكين بحكم الواقع، أو شريكة في التّضليل تغض الطرف عنه (Gillespie, 2017)، ويقترح «جيليبسي» تعريفًا لمفهوم المنصة (الإلكترونية) على النحو التالي: «هي بنية يمكن من خلالها التحدث أو التصرف، على غرار منصة قطار أو منبر سياسي». ومع ذلك، شأنه شأن عدد متزايد من الباحثين، يتجنب استعمال مفهوم «المنصات» لأنه يقلل من الدور الذي تلعبه الشركات المعنية. فيشير «جيليبسي» إلى أن المنصات الإلكترونية ليست في الواقع مساحات مسطحة ومفتوحةً وسالبةً بل «مشاهد متشابكة ومتعددة الطبقات تتمتع بملامح معقدة على سطحها وبجحور كثيفة في باطنها». وهذا يشير إلى أنّ شكل هذه البنية المعقدة يؤثر على كيفية نقل المُحتوى، حيث طرق النقل لا تكون دائماً مفتوحةً أو مباشرةً. فتتشكل طبيعة المُحتوى الذي يتلقاه مستخدمو الانترنت بفعل الخوارزميات، وقد تتغيّر أيضاً بشكل كبير بناءً على طلب من يتحكم بتصميم المنصة. وقد يسمح نموذج الأعمال لروبوتات الانترنت وللمستفزين bots and trolls بالتربص تحت السطح للانقضاض على الضحايا غير المتنبهين أو على أنواع محددة من المعلومات وفرض تحيزات معيّنة على غرار تقرير ما هو مسموح به أم لا وما هو الموضوع الأكثر انتشاراً على المنصة. وهذا هو أحد الأسباب الذي يفسر الاستخدام القليل لمصطلح «منصات» في هذا التقرير ليُستعاض عنه، حيث أمكن، بمصطلح «شركات التواصل عبر الإنترنت».

ومما يرتبط بهذه المسألة المكاسب المالية التي يتم تحقيقها من خلال تحليل كميات هائلة من البيانات المتاحة للشركات التي تسمح بنقل المُحتوى أو اكتشافه. وقد درست «زوبوف Zuboff, 2019» كيف يتم

....

[26] https://secure.avaaz.org/campaign/en/disinfo_hub/

إشـراك المسـتخدمين فـي إنتـاج هـذه البيانـات التـي يتـم بعـد ذلـك تسـخيرها لجنـي الأمـوال فيتـم بيعهـا لقدرتهـا علـى تحديـد مـا تسـميه زوبـوف «المسـتقبل السـلوكي». وتسـلّط تقاريـر منظمـة Ranking Digital Rights الضـوء علـى أن نمـوذج الأعمـال هـذا يـؤدي إلـى انتشـار أنـواع معينـة مـن المُحتـوى بمـا فـي ذلـك المعلومـات المُضلّلـة. فمـن خـلال إعطـاء الأولويـة لهـذا النـوع مـن المُحتـوى والتوصيـة بمحتويـات مماثلـة، يصبـح التّضليـل مرتبطـاً بشـكل متزايـد بإيـرادات المنصـات ومـزودي المُحتـوى، فتصبـح المشـكلة دائريـة (Marechal & Biddle, 2020, Marechal et al., 2020).

ويوفر كتاب «مجتمع المنصة: القيم العامة في عالم مترابط» أو «The platform society: Public values in a connective world» تحليلاً معمقاً لدور هذه الشركات في تحديد شكل المجتمع الحديث (Van Dijck et al., 2018) وهو يركز على القيم العامة في عالم يتم فيه التفاعل الاجتماعي بشكل متزايد على المنصات الرقمية، فيبحث في كيفية حماية هذه القيم. وحتى وقت قريب، كانت معظم الشركات تميل إلى التهرب من التزاماتها الاجتماعية الناجمة عن دورها كوسيط للمُحتوى، على الرغم من أن هذا قد بدأ يتغيّر بفعل الضغط الذي تمارسه عليها السُّلطات، ولا سيما صناع السياسات الأوروبيين. وعلى الرغم من أنّ بعض الشركات قد شجعت البحوث التي تركز على المعلومات المُضلّلة فهي تحجم عن إتاحة بياناتها لهذا الغرض. فعلى سبيل المثال، أعلنت شركة فيس بوك عن منح مليوني دولار أميركي للبحوث في «التّضليل والاستقطاب» شرط «ألا تعطى أي بيانات من فيس بوك وماسنجر وإنستغرام وواتساب وغيرها للفائزين بالمنحة».

ومـن المجـالات الأخـرى التـي تتطلّـب المزيـد مـن البحـث فـي مـا يخـص دور شـركات التواصـل عبـر الإنترنـت فـي مُكافحـة المعلومـات المُضلّلـة هـو اسـتغلال «الفراغـات البيانيـة» (Golebiewski & Boyd, 2019) وقد تمكنت الأبحـاث التـي كانـت تُجـرى خـلال كتابـة هـذا التقريـر البحثـي، كجزء من شـراكة بين فريق First Draft وباحثين من جامعة شيفيلد، من تحديد المشكلة التي تطرحها الفراغات البيانية خلال جائحة فيروس كورونا. فتبيّن أنـه عنـد قيـام الأفـراد بالبحـث عـن أجوبـة علـى أسـئلة حول أسـباب وأعراض وعلاج فيـروس كورونا، ظهر فراغ ناجـم عـن عـدم وجـود أجوبـة قابلـة للتحقـق (وذلـك يعود جزئيـاً إلـى عدم اليقين العلمي الفعلـي المرتبط بظهور فيـروس جديد؛ وفـي بعض الأحيـان بسـبب مـا تقـوم به السُّـلطات من كشـف انتقائي عن البيانـات الإحصائية). وقـد سـمح ذلـك لـوكلاء التّضليل باسـتغلال الفـراغ وسـده بمُحتـوى زائـف: «عندمـا يكـون هنـاك تكهنـات أو معلومـات مضللـة أكثـر مـن الحقائـق الموثـق بهـا حـول هـذه المصطلحـات، تقـوم محـركات البحـث غالبـاً بعـرض ذلـك علـى الأفـراد الذيـن قـد يكونـون يعيشـون لحظـة يـأس فـي خضـم الجائحـة. وهـذا مـن شـأنه أن يـؤدي إلـى الارتبـاك، ونشـوء نظريـات المؤامـرة، والمـداواة الذاتيـة، وتخزيـن المـواد، وتنـاول الجرعـات الزائـدة». (Shane, 2020) وبنـاءً علـى النتائـج الأوليـة، وإدراكاً للـدور الـذي تلعبـه مواقـع التواصـل الاجتماعـي كمحـركات بحـث بحكم الواقع، دعـا الباحثـون إلـى تطويـر أداة مشـابهة «لاتجاهـات غوغل» أو Google Trends ليصار إلى تطبيقها على مجموعـة مـن مواقـع التواصـل الاجتماعـي بمـا فـي ذلـك فيـس بـوك وتويتـر وإنسـتغرام ورديـت، بهـدف التوصـل إلـى تحديـد المعلومـات المُضلّلـة التـي تظهـر بفعـل أنشـطة البحـث هـذه بشـكل أسـهل وأكثـر شـفافيةً.

وقـد اجتـذب التقاطـع بيـن شـركات الإنترنـت وشـركات الإعـلام الإخباريـة كنواقـل للمُحتـوى الـكاذب بعـض التحليـلات. ويسـلط هـذا الضـوء بشـكل خـاص علـى التوتـر القائـم بيـن الصحافـة وشـركات التواصـل عبـر الإنترنـت فـي مـا يتعلق بجهود تنظيـم المُحتوى الراميـة إلـى مُكافحـة المعلومـات المُضلّلـة ونشـرها علـى نطاق واسـع والتـي غالبـاً مـا يسـتهدف موزعوهـا الصحفييـن والجهـات الناشـرة للأخبـار. وتعـود هـذه التوتـرات بجذورهـا إلـى علاقـة «الأصدقـاء-الأعـداء» بيـن هـذه الشـركات والجهـات الإخباريـة (Ressa, 2019)، والتـي

تفاقمت بسبب انهيار نماذج الأعمال الإخبارية التقليدية، وتآكل الأدوار الرقابية التاريخية أو ما يعرف بحراسة البوابة، وصعود «قوة المنصات» (Bell & Owen, 2017).

وقد أدى تصاعد المعلومات المُضلّلة الرقمية مع تزايد اعتماد الصحافة على شبكات التواصل الاجتماعي هذه لنشر المُحتوى والإشراك فيه، وتشجيع المنصات لهذا الاعتماد، إلى ظاهرة تعرف بعبارة «الاستيلاء على المنصة». ومن الأمثلة الأخرى على «الاستيلاء على المنصة» الطرق التي يمكن أن تؤدي بها الجهود المبذولة للحد من التّضليل إلى نتائج عكسية، كما حدث عندما قام واتساب بتغيير شروط الاستخدام في عام 2019 فأثر ذلك سلباً على قدرة وسائل الإعلام على استخدام التكنولوجيا لمواجهة التّضليل (Posetti et al., 2019b).

وتلتزم الصحافة التقليدية بمجموعة من القيم الإخبارية (Galtung and Ruge, 1965) تشمل الدقة والتدقيق في الخبر وتحقيق المصلحة العامة. ولكن قد لا تلتقي هذه القيم الا في نقاط قليلة مع القيم التي تلتزم بها المنصات الرقمية التي تشمل عادةً الابتكار والتواصل بين النظراء (Wilding et al., 2018)، ناهيك عن تحقيق الدخل على حساب المعايير التحريرية. وكما يُشير فوير (Foer, 2017)، فإن اعتماد وسائل الإعلام الإخبارية على قيم المنصات الرقمية، يعني أن سعيها المكثف نحو الانتشار قد يطغى على سعيها للبحث عن الحقيقة. وتتفاقم هذه المشكلة من خلال الخوارزميات التي تسعى إلى تحسين ونشر وحتى إنتاج الأخبار (Wilding et al., 2018) وتحسين نتائج البحث[27]. بالإضافة إلى ذلك، أصبح إشراك الجمهور محركاً أساسياً لوسائل الإعلام، مما أدى إلى تعديل إنتاج الأخبار لتصبح «أخف» (Hanusch, 2017) أي أقصر وأكثر اعتماداً على الصور وأكثر إثارةً للعواطف (Kalogeropoulos et al., 2016). إضافةً إلى ذلك، تقوم «مزارع المُحتوى» بإنتاج أو إعادة تدوير محتويات متدنية الجودة مشكوك في صحتها فيتم تحسينها لتحقيق إشراك الجمهور.

ولا شكّ في أنّ التحول الرقمي للصحافة مستمر حيث يُعتبر التغيير حالياً حالةً دائمةً. لذلك، من المهم أن يواكب البحث التحديات والفرص المرتبطة بإنتاج ونشر وتضخيم المعلومات المُضلّلة في النظام البيئي للأخبار في القرن الحادي والعشرين (Ireton & Posetti, 2018).

ويعرض القسم 6.1 من هذا التقرير تقييماً للإنترنت ولوسائل الإعلام الإخبارية الناقلة والعلاقة بينهما بالتفصيل.

4.2.3 المقاربات السياسية لدراسة التّضليل

لقد أفضت جائحة فيروس كورونا إلى مجموعة من الدراسات تهدف إلى تطوير استجابات سياسية، بما في ذلك من قبل اليونسكو (Posetti & Bontcheva, 2020a and 2020b) ومنظمة التعاون والتنمية في الميدان الاقتصادي (OECD, 2020) واستخدمت دراسة منظمة التعاون والتنمية في الميدان الاقتصادي إطار Wardle و Derakhshan (2017) لتحديد أربع استجابات للتضليل مرتبطة بالحوكمة: تحديد المعلومات وكشف زيفها؛ والمبادرات المدنية والإعلامية؛ واستراتيجيات التواصل؛ والتدابير التنظيمية. وتم التركيز بشكل خاص على التواصل العام وعلى الرسالة التي مفادها أنّه «على التواصل الاستراتيجي والشفاف أن يكون في خطوط العمل الأولى للمؤسسات العامة على جميع المستويات».

[27] https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2020/08/Follow-the-Money-3-Aug.pdf

ويشـرح تقرير «مدرسـة لندن للاقتصاد» بشـأن «مُعالجة أزمة المعلومات» (LSE, 2018) أو بالإنكليزية Tackling the Information Crisis كيـف أدت التغييـرات فـي النظـام الإعلامـي فـي المملكـة المتحـدة إلـى مـا يسـمى بأزمـة المعلومـات. ويؤكـد التقريـر أن ذلـك يتجلـى فـي انتشـار «خمسـة شـرور هائلـة» فـي صفـوف الجمهـور فـي المملكـة المتحـدة ألا وهـي: الارتبـاك والتشـكيك والتشـرذم وعـدم المسـؤولية واللامبـالاة. كمـا ويلخـص عـدداً مـن الاسـتجابات السياسـتية المتبعـة فـي المملكـة المتحـدة بمـا فـي ذلـك تحقيقـات برلمانيـة ومبـادرات حكوميـة تتضمـن علـى سـبيل المثـال Digital Charter أو ميثـاق رقمـي (UK DCMS & Rt Hon Matt Hancock, 2018b) وهـو ورقـة بيضـاء عـن القوانيـن الجديـدة تهـدف إلـى جعـل وسـائل الإعـلام الاجتماعيـة أكثـر أمانـاً (UK DCMS, Home Office, Rt Hon Matt Hancock & Rt Hon SajidJavid, 2018a)، ومختبـر الـذكاء الاصطناعـي الجديـد التابـع لمختبـر علـوم وتكنولوجيـا الدفـاع فـي بورتـان داون، والـذي يشـمل اختصاصـه «مواجهة الأخبـار الزائفة» (UK MOD et al., 2018). ومـن الاسـتجابات السياسـتية الأخـرى اسـتجابات مؤسسـية علـى غـرار مبـادرة الجهـة الناظمـة للاتصـالات فـي المملكـة المتحـدة Ofcom (2018b) ولجنـة الأخبـار الزائفـة، وتعليـم المهـارات النقديـة الأساسـية فـي المـدارس (National Literacy Trust, 2018). وفـي حيـن أن التقريـر يسـتعرض الاسـتجابات السياسـتية للمعلومـات المُضلّلـة بالتفصيـل، فإنـه يركّـز فـي المقـام الأول علـى التوصيـات والمبـادرات الحديثـة. لذلـك لا تـزال الحاجـة قائمـةً لإجـراء بحـوث حـول تحليـل نتائـج هـذه المبـادرات وتأثيرهـا. وتهـدف «مجموعـة عمـل حـرب المعلومـات» بجامعـة سـتانفورد[28]، التـي تـمّ إطلاقهـا فـي تشـرين الثانـي/نوفمبـر 2018، والتـي تتألـف مـن باحثيـن مـن مختلـف التخصصـات فـي مركـز الأمـن والتعـاون الدولييـن فـي «معهـد فريمـان سـبوغلي» و «مؤسسـة هوفـر»؛ إلـى «الدفـع قُدمـاً بفهمنـا للجوانـب النفسـية والتنظيميـة والقانونيـة والتقنيـة المتعلقـة بأمـن المعلومـات فـي حـرب المعلومـات»، والعمـل علـى وضـع مجموعـة مـن التوصيـات السياسـتية لمواجهـة تهديـدات المعلومـات المُضلّلـة الأجنبيـة. وقـد أنتجـت هـذا المجموعـة حتـى الآن عـدداً مـن الأوراق البيضـاء والتقاريـر. ويتضمّـن عملهـا بحوثـاً مـن تخصصـات مختلفـة، فيمـا تركـز فـي الوقـت نفسـه بشـكل ضيـق إلـى حـد مـا علـى الجوانـب السياسـية للتضليـل فـي الولايـات المتحـدة.

ومـن المـوارد الهامـة الأخـرى علـى المسـتوى السياسـي الأوروبـي؛ دراسـة أُجريت بتكليـف مـن دائـرة البحـوث البرلمانيـة الأوروبيـة للتحقيـق فـي آثـار مبـادرات مُكافحـة المعلومـات المُضلّلـة علـى حُريّـة التّعبيـر والتعدُّديـة الإعلاميّـة (Marsden & Meyer, 2019) إضافـةً إلـى تقريـر فريـق الخبـراء رفيـع المسـتوى حـول «الأخبـار الزائفـة» والمعلومـات المُضلّلـة علـى الإنترنـت (Buning et al., 2018).

ويبحـث التقريـر الأوّل فـي التنـازلات التـي يفرضهـا تطبيـق التقنيـات الآليـة لمواجهـة المعلومـات المُضلّلـة، مـع التركيـز بشـكل أساسـي علـى الطـرق التـي تسـمح باسـتخدام تشـريعات الاتحـاد الأوروبـي للدفـع بتصميـم هـذه التكنولوجيـات بطريقـة لا تُقيّـد حُريّـة التّعبيـر مـن دون سـبب، وتُعـزز فـي الوقـت نفسـه الشـفافية والمُسـاءلة إلـى أقصـى حـد. ويركّـز التقريـر فـي المقـام الأول علـى الاسـتجابات التكنولوجيـة والتشـريعية للمعلومـات المُضلّلـة، ويثيـر مخـاوف حـول طبيعـة التشـريعات الحاليـة التـي قـد تقيّـد حُريّـة التّعبيـر؛ فيخلُـص إلـى أن هنـاك نقـص فـي البحـوث السياسـتية فـي هـذا المجـال، وأنّ الحلـول الوحيـدة لا سـيما تلـك التـي تُركّـز علـى الاسـتجابات التكنولوجيـة بشـكل أساسـي غيـر كافيـة. وعلـى نفـس المنـوال، يقـدم تقريـر فريـق الخبـراء رفيـع المسـتوى منظـوراً مبنيـاً علـى السياسـات وينصـح بعـدم اتبـاع حلـول سـطحيّة بل حلـول شـاملة تُعزّز من الشـفافية والدراية المعلوماتيـة والتمكيـن إلـى أقصـى حـد، كمـا ويقتـرح مزيجـاً مـن الإجـراءات قصيـرة وطويلـة الأجـل.

[28] https://cisac.fsi.stanford.edu/content/information-warfare-working-group

وهكذا، يركز هذان التقريران على وجه التحديد على قضايا السياسات الأوروبية، ولكن لا يذهبان أبعد من ذلك؛ فلا ينظران في كيفية تطبيق هذه الأمور في مناطق أخرى. والحقيقة أن في جميع التقارير الحالية المبنية على السياسات ثغرةً بحثيةً كبيرةً تتمثل في أنّ كل تقرير يقترح مجموعةً من الاستراتيجيات الخاصة به، ولكن لا يُوضّح كيفية الانتقال من الاستراتيجيات إلى استجابات أكثر شمولية، على الرغم من أنّ المعلومات المُضلّلة لا تحترم الحدود الجغرافية ولا الحدود السياسية.

من ناحية أخرى، أصدر فريق من الخبراء من «جامعة بنسلفانيا» تقريراً بعنوان «الحُريّة والمُساءلة: إطار عابر للأطلسي لتنظيم الخطاب على الإنترنت» (Annenberg Public Policy Center, 2020). وتنص هذه الوثيقة المعنونة بالإنكليزية Freedom and accountability. A transatlantic framework for moderating speech online" على التالي: «من منظور حُريّة التّعبير، قمنا بتحليل مجموعة من الممارسات المعتمدة من قبل الشركات كما نظرنا في القوانين والمقترحات السياسية وقمنا بجمع أفضل الممارسات لتقديم إسهامات مدروسة في إطار المناقشات الدائرة حول الإطار التنظيمي في أوروبا وأمريكا الشمالية». ولمعالجة المشكلات على شبكة الإنترنت، بما في ذلك المعلومات المُضلّلة، يقترح التقرير أن تعمد الدول إلى تنظيم شركات الإنترنت على أساس أحكام الشفافية الإلزامية، وأن يكون هناك أيضاً رقابة ناظمة تسعى إلى إلزام المنصات «بالوفاء بوعودها». وبالنسبة لشركات الإنترنت نفسها. يقترح التقرير هيكلاً للإفصاح من ثلاثة مستويات، وآليات إنصاف فعالة، وإعطاء الأولوية لمعالجة سلوك «الفاعلين الأشرار» على الإنترنت قبل معالجة المُحتوى نفسه.

5.2.3 الدراسات والموارد ذات الصلة بالممارسة

يوثق تقرير الأخبار الرقمية[29] Digital News Report الذي يشمل دراسةً لـ40 سوقاً والمُعد من قبل «معهد رويترز لدراسة الصحافة» الدور الذي تقوم به شركات الإنترنت في مجال نشر المعلومات وما يسميه مؤلفو التقرير «المعلومات المغلوطة». ويشير إلى أنّ الجماهير يمكن أن «تصل أيضاً إلى معلومات مضللة (كما تصل إلى غيرها من المعلومات) عن جنب أي من خلال محركات البحث، ووسائل التواصل الاجتماعي، أو غيرها من أشكال الاكتشاف الموزع». وقد قام تقرير عام 2018 بدراسة الاختلافات في التعرض والمخاوف، والمعتقدات المختلفة حول علاجات المُحتوى الكاذب على الإنترنت. كما وبحث المعهد في أنواع المعلومات المغلوطة المتعلقة بجائحة فيروس كورونا ومصادرها والادعاءات بخصوصها (Brennan et al., 2020)، وقام بمسح الاستجابات للمعلومات المُضلّلة من ثلاث وكالات إخبارية في دول الجنوب تعرّف عن نفسها بأنها «في الخطوط الدفاعية الأمامية» في «حرب التّضليل» (Posetti et al., 2019a; Posetti et al., 2019b). وتحدد هذه التقارير الأخيرة أساليب مُطوّرة لإعداد التقارير الاستقصائية (بما في ذلك تحليل البيانات الضخمة والشبكات)، وتقنيات متطورة لإشراك الجمهور (مثل: الاستجابات التعاونية لإظهار وكشف المعلومات المُضلّلة)، ونماذج صحافية تعتمد «المناصرة» أو «النشاط النضالي» (تنطوي على تنظيم حملات ضد نواقل المعلومات أو توفير التدريب على الدراية بوسائل الإعلام الرقمية للمجتمعات المحلية) كوسيلة للاستجابة لأزمة المعلومات المُضلّلة.

[29] http://www.digitalnewsreport.org/

وقد أنتجت «وحدة جيكسو» (وهي ذراع من غوغل) ما تسميه صورةً لحملات التّضليل في جميع أنحاء العالم، لدعم نظريتها التي مفادها أن «فهم عمل حملات التّضليل هو من الخُطوات الأُولى التي تسمح بمكافحتها على الإنترنت»[30]. وتؤكد «جيكسو» أنّ هذه الصورة تستند إلى بحث أجراه «المختبر الجنائي الرقمي التابع للمجلس الأطلسي» ويعكس نظرته في التوصيف. بالإضافة إلى ذلك، يذكر الباحثون في «جيكسو» أنّ تصورهم يستند في المقام الأول إلى تقارير صحافية مفتوحة المصدر باللُّغة الإنكليزية عن الحملات المستهدفة للغرب. وهذه الأنواع من الصور توفر نظرةً عامةً مثيرةً للاهتمام، على الرغم من القيود الجغرافية، ولكنها قد تخلط أنواعاً مختلفةً جداً من المعلومات المُضلّلة.

وتشمل مجموعة جديرة بالذكر من الموارد العملية المتعلقة بالمعلومات المُضلّلة مواد ستناقش في القسمين 7.1 (الاستجابات المعيارية والأخلاقية) و7.3 من الفصل السابع (حول الاستجابات التعليمية)، لا تدعم المهارات العملية فحسب، بل تبحث أيضاً في النظريات والاتجاهات الكامنة. وإنّ دليل اليونسكو «الصحافة، والأخبار الزائفة والتّضليل» (Ireton & Posetti, 2018) هو مورد تعليمي قائم على البحوث يستهدف الصحفيين والوكالات الإخبارية التي تواجه المعلومات المُضلّلة مع التركيز على قضايا حُريّة التّعبير. بالإضافة إلى دور الدليل كمجموعة من الموارد الداعمة للتدريب الصحفي، فإنه يبحث أيضاً في طبيعة الصحافة لفائدة الثقة، والتحديات الهيكلية التي سمحت بذيوع صيت المعلومات المُضلّلة وانتشارها الهائل، والقنوات التي تؤدي إلى اضطراب المعلومات مثل التكنولوجيا الرقمية ووسائل التواصل الاجتماعي. كما ويصف الدليل عملية استهداف الصحفيين ومصادرهم في سياق حملات التّضليل. ويقدم إطاراً يسعى إلى فهم الصحافة المستقلة والنقدية كآلية لمُكافحة المعلومات المُضلّلة. كما ويوفر نماذج للاستجابة بشكل مبتكر للتحديات التي تطرحها المعلومات المُضلّلة لفائدة تأثيرها على الصحافة والجمهور. ومن بين الموارد الأخرى من هذا النوع دليل التحقُّق من المعلومات المُضلّلة والتلاعب الإعلامي الصادر عن مركز الصحافة الأوروبي (Silverman, 2020).

ومن الأمثلة على الموارد التي تركز على الجمهور الأداة الرقمية حول الدراية الإعلامية والمعلوماتية الصادرة عن اليونسكو[31] والمؤلفة من وحدات تعليمية كبيرة حول الدراية الإعلامية والمعلوماتية بعدة لغات، والوحدة التعليمية الصادرة عن المركز الدولي للصحافيين حول تاريخ التّضليل (Posetti & Matthews, 2018). ومن الثغرات الهامة في عدد من هذه الأدوات والبرامج هو التركيز على التمثيل الأوسع «للبيانات»، بما في ذلك قضايا الخصوصية والتنميط، وبشكل أشمل كيفية جمع البيانات واستخدامها من قبل المنصات الإلكترونية، كما نوقش سابقاً في هذا الفصل. فقد أظهرت تجربة منظمة 5Rights غير الحكومية، على سبيل المثال، أنه عندما يُدرك الأطفال هذه المفاهيم، فإن درايتهم بالمعلومات تتحسّن بشكل عام أيضاً. ومع ذلك، فإن العديد من المقاربات القائمة على المهارات والهادفة إلى التصدي للتضليل لا تُركّز إلا على المفاهيم الأولية للتحقُّق من المصادر المباشرة من دون النظر في هذه الجوانب التأسيسية الأوسع نطاقاً[32].

وأخيراً، قامت «الوكالة السويدية للطوارئ المدنية» بإصدار دليل مُوجَّه إلى الجهات الحكومية العاملة في مجال التواصل حول مُكافحة أنشطة «تأثير المعلومات» (MSB, 2020).

30 https://jigsaw.google.com/the-current/disinformation/dataviz/

31 https://en.unesco.org/MILCLICKS

32 https://5rightsfoundation.com/uploads/digital-childhood---final-report.pdf

3.3 الثغرات البحثية الحالية

كما ذُكر سابقاً، هناك عدد كبير من البحوث المتعلقة بالتّضليل وبالاستجابات المتبعة في مواجهته من الناحيتين النظرية والعملية على حدٍّ سواء. ومع ذلك، يبدو أن هناك فجوة واضحة في البحوث الأكاديمية والتحقيقات الصحفية والدراسات التي تتم بتكليف من منظمات المجتمع المدني والمنظمات الحكومية الدولية. وبالإضافة إلى ذلك، يبدو أن التعاون الفعلي بين هذه القطاعات نادر الحدوث. وقد وُضعت هذه المبادرات والمنشورات المذكورة وفقاً للطلب، وهي تغطي مواقع مختلفة مما يجعل من الصعب تعقبها وتحليلها وتوليفها بطريقة متسقة. فعلى سبيل المثال، لا تزال الدراسات القائمة بين عدة مؤسسات حول العلاقة بين المنطق التكنولوجي/التجاري ومجال سياسات الشركات والدولة ضعيفة، كما سيُناقش ذلك بمزيد من التفصيل في القسم 6.2 من الفصل السادس.

كذلك، لم تتم دراسة أثر معظم الاستجابات الهادفة إلى التصدي للتضليل بشكل كاف. ففي حين أنّ بعض الأبحاث قد نظر إلى تصديق المجموعات (مثل كبار السن) بشكل خاص للمعلومات المُضلّلة ومشاركتها (Carey et al., 2020; Guess et al., 2019)، لم يكن هناك سوى عدد قليل من الاستجابات التي تستهدف هذه الفئات الضعيفة. كما ونلاحظ ندرةً في التقييم التجريبي لهذه الاستجابات مع استثناءات بسيطة كما في حال دراسة Humprecht et al أو همبرخت وآخرون (2020) وإن كانت محدودةً في تركيزها الجغرافي. كذلك، فإن أساليب التصدي للتضليل لم تغطي بما فيه الكفاية مفهومي الولاء الجماعي وانعدام الثقة في السلطة اللذين يتطلبان نظرةً مختلفةً ومعالجةً خاصةً لمسائل أكثر جوهرية.

وأخيراً، على الرغم من وجود مجموعة متنامية من البحوث والبرمجيات وتطوير للتدريب والموارد التي تركز على التصدي للمعلومات المُضلّلة، فإن هناك غياب للبحوث التي تركز على المعلومات المُضلّلة من منظور حقوق الإنسان وحُريّة التّعبير والنفاذ إلى تكنولوجيا النطاق الواسع واستخدامها في جميع أنحاء العالم رغم أنها لا زالت في بداياتها.

وفي ما يلي تحليل إضافي يغطي مجالات محددة اتضح وجود ثغرات هامة فيها:

معالجة الفروقات والصلات بين مجالات المعلومات المُضلّلة المختلفة

من حيث الأطر العامة، لا تُميّز معظم البحوث المنشورة بوضوح بين الأنواع الجديدة من المعلومات المُضلّلة (على سبيل المثال، التزييف الفج) وتلك الأقدم منها بشكل ملحوظ (على غرار تأثير المعلومات الذي يتداخل مع التّضليل كما نوقش أعلاه). ولا ينطبق البعض الآخر إلا على سياقات محددة، مثل التّضليل السياسي، أو قد يكون غير قابل للتطبيق إلا في الدول غير الغربية (Brooking et al., 2020). وينظر عدد من الأطر أيضاً إلى المعلومات المُضلّلة ليس فقط من منظور سياسي، بل من ناحية النفوذ الأجنبي في المقام الأول. وبالتالي لا تعالج هذه الأطر القضايا العديدة المتعلقة بالمعلومات المُضلّلة المحلية، مثل تلك المتعلقة بالأزمات الصحية، وقضايا الهجرة، والتواصل في حالات الكوارث.

من ناحية أُخرى، أُجريت دراسات منفصلة حول هذه القضايا، كما تشهد على ذلك الجهود الأخيرة الساعية إلى مسح المعلومات المُضلّلة حول جائحة فيروس كورونا وتنفيذ استراتيجيات مضادة لها،

والتي ستتم مناقشتها بمزيد من التفصيل في الفصول التالية. بشكل عام، أدى الوباء إلى زيادة الوعي العام حول المعلومات المُضلّلة وإلى تعزيز الجهود التثقيفية من قِبل الجهات الفاعلة الحكومية وغير الحكومية (الحكومات، وشركات التواصل عبر الإنترنت، والشركات الإعلامية، وما إلى ذلك). وقد شكّلت جائحة فيروس كورونا مثالاً واضحاً عن آثار المعلومات المُضلّلة ومضارها، وبالتالي رفعت من أهمية هذه الظاهرة في نظر الجمهور، وقد تؤدي إلى تزايد البحوث في هذا الخصوص، على غِرار «مبادرة منظمة الصحة العالمية» للتعمق في دراسة المعلومات عن علم الأوبئة أو «Infodemiology» وهو مجال متعدد التخصصات له صلة بمجالات خارج مجال الصحة[33].

توافر البيانات للبحوث

في ما يتعلق بفهم طبيعة المعلومات المُضلّلة ونشرها والأنشطة الهادفة إلى مكافحتها، تشكل مسألة عدم شفافية الخوارزميات المستخدمة في منصات التواصل الاجتماعي والقضايا المتعلقة بالنفاذ إلى بياناتها عائقاً خطيراً، كما هو وارد في القسم 4.2 من الفصل 4. ويعتمد القياس الكمي للمعلومات المُضلّلة على الإنترنت على الافصاح الانتقائي من قبل الشركات وما يرد في تقارير الشفافية التي تصدرها، وذلك من دون أن يتمكن الباحثون من النفاذ إلى البيانات الأصلية.

وتشير أدلة من دراسات خارجية إلى وجود حالات تضليل تم اكتشافها من خلال تقنيات تحليل المُحتوى. وقد أكدت إحدى الدراسات المقتضبة أن مقطع فيديو واحد من أصل كل أربعة مقاطع تتمتع بالشعبية على يوتيوب حول فيروس كورونا يحتوي على معلومات مُضلّلة[34]. وقد حلّل هذا التقرير البحثي 69 مقطعاً من أكثر مقاطع الفيديو مشاهدةً باللغة الإنكليزية من يوم واحد في آذار/مارس 2020 فوجد أنّ 19 منها يحتوي على معلومات لا تعتمد على وقائع ولكنه حظي بأكثر من 62 مليون مشاهدة. وفي دراسة أخرى قامت بتحليل أكثر من 1300 صفحة على فيس بوك لديها حوالي 100 مليون متابع تبيّن أنه في حين أن الصفحات المناهضة للتلقيح لديها عدد أقل من المتابعين من الصفحات المؤيدة للقاح، إلا أنها أكثر تجمعاً وأسرع نمواً، ومترابطة على نحو متزايد بصفحات أخرى[35].

وتؤكد هذه النتائج أهمية تقييم أنماط التّضليل على الإنترنت، كما أنها تظهر ما يمكن القيام به حتى من دون الإفصاح عن البيانات من قبل شركات الإنترنت.

ومع ذلك، فإنّ معظم البحوث المتعلقة بالتّضليل محدودة بفعل عدم نفاذها إلى مجموعات البيانات الكاملة من شركات التواصل عبر الإنترنت. وهذا يؤدي إلى غياب العمق في التحليلات التي تُقيّمها. كذلك تتم الدراسات عادةً حول منصة مختارة (عادةً تويتر، بفعل حجم بياناته المفتوحة والعامة)، لا حول عدة منصات. ونادراً ما يتم النظر في تطبيقات المراسلة بسبب طبيعتها المغلقة. وتطرح شركات التواصل الاجتماعي عقبات أمام البحوث المستقلة حيث أنها تقطع طريق الوصول إلى واجهات برمجة

....

33 https://www.who.int/docs/default-source/epi-win/infodemic-management/infodemiologyscientific-conference-booklet.pdf?sfvrsn=179de76a_4

34 https://www.nbcnews.com/health/health-news/live-blog/2020-05-13-coronavirusnews-n1205916/ncrd1206486#liveBlog-Header

35 https://www.nature.com/articles/s41586-020-2281-1

التطبيقات APIs التي يمكن للباحثين من خلالها جمع البيانات ذات الصلة، مما يعكس إلى حد ما المشكلات المتعلقة بمحركات البحث، إذا لا يستطيع سوى من له علاقات مباشرة مع شركات البحث الكبرى العمل بفعالية (Walker et al., 2019). على سبيل المثال، من الصعب معرفة التفاصيل عندما تتم إزالة المستخدمين أو الرسائل من قبل المزود (أو عندما يسحب المستخدم المعلومات من تلقاء نفسه). وفي حين أن هذه المنصات تتيح لمجموعة مختارة من الباحثين الأكاديميين النفاذ إلى هذه البيانات عن طريق المنح البحثية[36] عن طريق أدوات مثل Crowdtangle[37]، كانت تلك الإمكانية محدودة النطاق عند كتابة هذا التقرير وتتضمن قيوداً على أنواع البحوث التي يمكن القيام بها. وفي ضوء جائحة فيروس كورونا، أطلقت Crowdtangle (في آذار/مارس 2020) أكثر من 100 منصة LiveDisplays متاحة للجمهور تسمح للباحثين التحقيق في قضايا مثل: انتشار المعلومات حول الجائحة على وسائل التواصل الاجتماعي. وعلى الرغم من هذه الإمكانية؛ لا يوفر ذلك إلا مجموعةً محدودةً من البيانات.

ويناقش تقرير «مركز أننبرغ للسياسات العامة» (2020)، المشار إليه أعلاه، بأن الشفافية تمكّن الحكومات من وضع سياسات قائمة على الأدلة للإشراف على شركات الإنترنت، وتدفع بالشركات إلى دراسة المشكلات التي لم تكن لتعالجها لولا ذلك، وبالتالي تقوم بتمكين المواطنين. ويلقي هذا الضوء على أهمية قيام الشركات بضمان نفاذ أوسع بكثير إلى البيانات. ومن المفهوم أن تكون الشركات حساسةً إزاء توفير البيانات، وذلك لأسباب تتعلق بالسرية التجارية، كما وأنها تتجنب التنازل عن البيانات كما حدث خلال مشكلة Cambridge Analytica. على هذه الخلفية، اقترح «ماكارثي» (2020) أنموذجاً متدرجاً للوصول إلى بيانات الشركات، يميّز بين المستويات المختلفة التي يمكن أن تتاح للجمهور والباحثين المصادقين والجهات الناظمة.

استخدام المعلومات المُضلّلة والاستجابة لها

لا تزال الدراسات المتعلقة بسلوك المستخدمين وتصورهم ناقصة، ولا سيما في ما يتعلق بالعلاقة بين آثار المعلومات المُضلّلة والأخبار. على سبيل المثال، حتى عندما تُعرض على الأفراد مجموعة متنوعة من المقالات، يميلون إلى اختيار المقالات الإخبارية الأكثر تماشياً مع مُعتقداتهم الخاصة (Kelly Garrett, 2009) من خلال الخصائص المبنية على تفضيلات المُستخدم أو التعرُّض الانتقائي، وهما أمران تعززهما الخوارزميات التنبؤية. ومع ذلك، لم يُنجز سوى القليل حول تقييم الأثر الفعلي لذلك. ويترتب على ذلك تداعيات هامة على التّضليل في ما يخص قضايا الدعاية أو المعتقدات الخطيرة المتصلة بالصحة، مثل تلك التي يروج لها مناهضو التلقيح. وتثير آثار هذا التعرض الانتقائي مخاوف متزايدة، لأنها قد تؤدي إلى تعزيز التشرذم الاجتماعي، كما وقد تعكس أو تُضخّم الانقسامات القديمة، ممّا يعزز أيضاً الآراء والتحيزات الإدراكية القائمة. كما يتطلّب الربط بين التعرض للتضليل وتأثيره على السلوك خارج الإنترنت مزيداً من التحقيق. وينطبق ذلك على العلاقة بين المعلومات المغلوطة والخوف والذعر والسلوك غير الأناني وغير العقلاني (Osmundsen et al., 2020).

....

36 https://about.fb.com/news/2019/04/election-research-grants/

37 https://www.crowdtangle.com/

وتوجد حالياً مفاهيم متنافسة حول مدى تأثير تعرض شخص ما لوجهات نظر مختلفة على وجهات نظره الأيديولوجية الخاصة. فمن ناحية، يعمل الاستخدام المتزايد لوسائل التواصل الاجتماعي والأخبار الشخصية كـ«فقاعة مصفاة» أو «غرفة صدى»، مما يعزز المعتقدات القائمة والفصل الأيديولوجي. ومع ذلك، هناك مجموعة متزايدة من البحوث التجريبية التي تؤكد أن تأثير غُرف الصدى مبالغ به (Dubois & Blank, 2018; Guess et al., 2018a)، وأن مجموعة فرعية صغيرة فقط من الناس تميل إلى استخدام وسائل الإعلام بطريقة مخلة بشكل كبير (Gentzkow & Shapiro, 2011; Flaxman et al., 2016). وهذا ما ينسحب أيضاً على المعلومات المغلوطة (Guess et al., 2018b). ويرى آخرون أن «تُخمة» المعلومات، إلى جانب ما يترتب عليها من تنوع أكبر للمعلومات المستهلكة، يوسع في الواقع من نطاق مصادر الأخبار التي يتعرّض لها الناس (Fletcher & Nielsen, 2018). وأظهرت دراسة أخرى أن المستخدمين من خلفيات سياسية مختلفة للغاية يتعرضون لمجموعات مماثلة للغاية من الأخبار السياسية (Nechushtai & Lewis, 2018) مما يناقض النظريات حول آثار غُرفة الصدى على وضع الخصائص الشخصية للأخبار.

وما يبقى غير واضح هو الأثر الناجم عن توسيع نطاق التعرُّض لمنظورات مختلفة على قضايا التحزب الأيديولوجي. إنّ فهم وقياس التنوع الأيديولوجي عبر البيانات الاجتماعية الضخمة، والتأثير على المنظورات الأيديولوجية التي قد تنشأ عن التعرض لهذا التنوع، من شأنهما أن يؤديا إلى تحسين فهم تأثير المعلومات المُضلّلة والمضمون المضاد لها على غِرار التحقُّق من الوقائع والتدقيق في الأخبار الصحفية. ولابد من إجراء دراسات واسعة النطاق حول المستخدمين لنفهم بطريقة أفضل كيف يقوم هؤلاء بتقييم حقيقة المعلومات وموثوقيتها سواءً من منظور عملي أو من منظور نفسي. كذلك، فإنّ الدراسات التي تستهدف سلوك المستخدمين في ما يخص التعامل مع المعلومات الموثوق بها والتحقُّق منها وإعادة نشرها، مثل تلك التي يصدرها ناشرو الأخبار والصحفيون المستقلون، يمكن أن توفر نظرةً ثاقبةً حول هذه الموضوع. وقد بُذلت عدة محاولات للحد من أثر التحيز في نظم المعلومات لدعم عملية تكوين الرأي من دون تأثير عوامل الانتقاء. وقد ركّز بعضها على جعل المستخدمين يُدركون وجود التحيز من خلال تنبيههم إلى ذلك (Epstein & Robertson, 2015)، ومن خلال تزويدهم بتصوُّر عن سلوكيات القراءة والتحيز (Munson et al., 2013)، أو الإشارة إلى صفحات الإنترنت التي تقدم آراءً مختلفةً عن الآراء القائمة. ويعتمد آخرون على التصورات لدعم التنوع في أنشطة الإنترنت (Graells-Garrido et al., 2016)، وأنظمة التوصيات (Tsai and Brusilovsky, 2018)، ونتائج البحث (Verberne, 2018). ويركز البعض على شفافية الخوارزميات من خلال شرح كيفية عمل المصفاة وتمكين المستخدم من التحكم في الخوارزميات وبالتالي غُرف الصدى الخاصة به (Nagulendra & Vassileva, 2014). ذلك، ويحاول آخرون كسر غُرف الصدى المحتملة من خلال تصميم البرمجيات وواجهات المستخدم (Bozdag, E., & van den Hoven, J., 2015). ومع ذلك، فإن نجاح هذه المقاربات جميعها محدود بعض الشيء (Faridani, 2010; Liao & Fu, 2013)، ومن الواضح أن هناك حاجة إلى مزيد من الدراسات لفهم أنماط وعادات استخدام الأخبار على الإنترنت بشكل أفضل، مثل كيفية تنقل الأفراد بين البيئات المتغيرة باستمرار لاختيار الأخبار التي يقررون قراءتها (Swart et al., 2017).

مشهد المعلومات التكنولوجية والمؤسسية المتغيّر

إنّ العديد من الاستجابات للمعلومات المُضلّلة الوارد وصفها في هذا التقرير لا تزال جديدةً ولم تنفذ بعد على نطاق واسع. وقد يرجع ذلك إلى أنّ التكنولوجيات لا تزال قيد التطوير أو التحسين، لأنها

غير معتمدة إلا من قبل أقلية صغيرة، أو لأسباب أخرى، كالمسائل القانونية والأخلاقية التي تحتاج إلى حل. فعلى سبيل المثال، عندما لا تُستخدم مقاربات المصداقية والتصنيف على نطاق واسع، يقوّض ذلك فعاليتها وفهم إمكاناتها. ولا يُعرف ما إذا كانت ستتجح هذه المقاربات إلا متى تطبّق على نطاق أوسع ولكن هناك أيضاً آثاراً خطيرةً محتملةً قد تنتج عن تطبيقها على نطاق واسع، كما هو مفصل في تقرير «Ticks or It Didn't Happen» الصادر عن منظمة Witness في المرجع التالي (Witness Media Lab, 2019). وهذا يوضح «معضلة كولينجريدج» (Collingridge, 1980)، التي تفترض أنه لا يمكن التنبؤ بالعواقب الاجتماعية للتكنولوجيا في كثير من الأحيان إلى أن يتم تطوير هذه التكنولوجيا بالفعل، وعندئذٍ يكون الأوان قد فات في معظم الأحيان، أو على الأقل يصبح التغيير أكثر صعوبةً. ولا يشير لا تقرير كولينجريدج ولا تقرير Witness إلى أنه لا يمكن التغلب على هذه التحديات، بل يُركّز التقريران على ضرورة النظر المُبكّر في السيناريوهات، واعتماد مقاربة مرنة للتعامل معها.

وفي الإطار نفسه، لا يزال هناك نقص في تقييم العديد من التكنولوجيات المقترحة لمواجهة المعلومات المُضلّلة، كما ولا يناقش ذلك إلا قليلاً. كذلك لا تُعرف دائماً فعالية بعض المنهجيات، كمفهوم التحقُّق من الوقائع، حيث أن البحوث قد أظهرت أن المواد التي يتم التحقُّق منها غالباً ما تختلف كثيراً عن المعلومات المُضلّلة من حيث انتشارها. والواقع أنه في بعض الحالات ظهر «أثر عكسي» لتصحيح موضوع ما؛ حيث أدى التصحيح إلى زيادة التصورات الخاطئة (Nyhan, 2012). ومن شأن إجراء المزيد من البحوث أن يساعد في تقييم أثر التكنولوجيات ونظريات التغيير الكامنة الخاصة بها، والتي قد تستند إلى افتراضات خاطئة أو مُضلّلة. ويرد مزيد من المناقشة حول هذا الموضوع في القسم 4.1 من الفصل الرابع.

تمثيل البحوث للمُستوى الدُّولي

تُعاني دول الجنوب بصفةٍ خاصة من نقص التمثيل في البحوث. ومن الأمثلة على ذلك «دراسة تشاتورفيدي» Chaturvedi للهند (Chaturvedi, 2016)؛ ودراسة «كاور وآخرين» Kaur et al لآسيا والمحيط الهادئ (Kaur et al., 2018)؛ والتقارير الحديثة للبحوث المشتركة بين مؤسسات FullFact وChequeado وAfricacheck التي ركزت على الأرجنتين وجنوب أفريقيا ونيجيريا[38]، وتقرير Oxtech حول مبادرات مُكافحة المعلومات المُضلّلة التي تستخدم أمثلةً من 19 بلداً في أربع قارات[39]. وتركز التقارير المبنية على منظور صانعي السياسات على أوروبا وأمريكا الشمالية بشكل حصري تقريباً. وهذه فجوة واضحة تسعى هذه الدراسة إلى معالجتها، حيث تأمل بأن يؤدي ذلك إلى الاستثمار في بحوث مُوجّهة نحو العمل في المستقبل.

ويسلّط كل هذا الضوء على قيمة دراسة عالمية واسعة النطاق على غرار هذه الدراسة التي تجمع بين تعدُّد أنواع الاستجابات للمعلومات المُضلّلة، والتي تعرض لاحتياجات وتحديات المناطق الجغرافية المغايرة ثقافياً.

38 https://fullfact.org/research/

39 https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/08/A-Report-of-AntiDisinformation-Initiatives

البعد المتعلّق بحقوق الإنسان

إنّ قلة من الأُطُر المفاهيمية أو غيرها من المؤلّفات تُركّز بشكل فعلي على المشكلة الأساسية المتمثلة في ضمان التوازن بين حماية حُريّة التّعبير ودعم مفاهيم الحقيقة مقابل التّضليل، على الرغم من أن هذا يرتبط ضمناً ببعض المناقشات في هذا الفصل حول شركات التواصل عبر الإنترنت، وحول النزاهة الصحفية. وفي الوقت نفسه، من الواضح أن تنظيم الخطاب على وسائل التواصل الاجتماعي في محاولة لمنع التّضليل؛ له تداعيات أخلاقية وسياسية تتقاطع مع حُريّة التّعبير، كما هو الحال بالنسبة لتشريعات حول «الأخبار الزائفة» التي تمثل تهديداً كبيراً على حُريّة الصحافة. وقد تم مؤخراً انتقاد «مدونة السُّلوك» الصادرة عن «الاتحاد الأوروبي» حول المعلومات المُضلّلة (European Commission, 2018c) لسماحها نظرياً، بل حتى لتحفيزها فرض قيود على حُريّة التّعبير يُزعَم أنها قانونية من الناحية الفنية (Kuczerawy, 2019). ويعبّر Kuczerawy عن مخاوفه من أن يؤدي تجنيد منصات خاصة للحد من محتويات معينة لا تعد غير قانونية؛ إلى عواقب غير مقصودة، فيؤكد بأنه من الصعب «تقييم تأثير المدونة على ممارسة الحق في حُريّة التّعبير». وفي الدول من خارج الاتحاد الأوروبي التي تعتمد لوائح تنظيميةً أقل صرامة، تزداد المخاوف إزاء هذا النوع من الممارسات. وتناقَش هذه المسائل بمزيد من التفصيل لاحقاً في هذا التقرير، ولا سيما في مناقشة الاستجابات التشريعية للمعلومات المُضلّلة الواردة في القسم 5.1 من الفصل الخامس، وكذلك في مناقشة الاستجابات السياستية الواردة في القسم 5.2 من الفصل الخامس، حيث أن هذين النوعين من الاستجابات يُعنيان بهذه المسألة.

4.3 الإسهامات الجديدة التي تُقدّمها هذه الدراسة

بعد وضع هذه الدراسة في سياق الأُطُر النظرية القائمة والبحوث السابقة، وبعد تحديد الثغرات في البحوث الحالية حول موضوع المعلومات المُضلّلة، يُسلّط هذا القسم الضوء على الإسهامات الجديدة التي تقدمها هذه الدراسة.

أولاً، سعت هذه الدراسة إلى اعتماد نظرة عالمية، في حين ركزت العديد من التقارير المذكورة أعلاه إلى حد كبير على بلدان أو قارات معينة، وركز قدر كبير من البحوث على حالات من المملكة المتحدة والولايات المتحدة و/أو أوروبا. وهذا يعكس جزئياً الحقيقة؛ وهي أنّ هذه المناطق الجغرافية تتمتع بنشاط عال في مجال التصدي للتضليل، كما وأنها المكان الذي تتركز فيه غالبية الأبحاث والتمويل حول هذه الظاهرة. كذلك، تقتصر مصادر التّضليل التي تجري دراستها في تقارير أخرى على المحتويات باللُّغة الإنكليزية.

وعلى النقيض من ذلك، يسعى هذا التقرير إلى التركيز على قضايا ومبادرات من جميع أنحاء العالم، بما في ذلك تلك الواردة من أفريقيا وأستراليا وأوروبا الوسطى والشرقية وأمريكا اللاتينية وآسيا. وقد ساعد ذلك في الكشف المُبكّر عن أن بعض الاستجابات الصحفية للتضليل تعتمد على وجود متطلبات تكنولوجية معينة، أو يصعب اعتمادها من قبل الذين يعيشون في حالات النزاع أو القلاقل (مثلاً، عندما يحتاج المراسلون إلى عدم الكشف عن مصادرهم فيتعذر نتيجةً لذلك عليهم استخدام أدوات نقطة الالتقاط point-of-capture للصُّور الفوتوغرافية ومقاطع الفيديو). وفي ما يلي، نناقش كيف ولماذا قد تكون استجابات معينة صعبة بالنسبة إلى الجهات الفاعلة في بعض البلدان والأوضاع، وهي مسائل لا ينظر فيها الفاعلون في أوروبا الغربية والولايات المتحدة.

كذلك، إنّ مؤلفي هذا التقرير هُم من خلفيات إثنية ومن مناطق متنوعة، وهم يتكلّمون لغات مُختلفة، ولديهم معرفة بالحالات من مختلف أنحاء العالم. وهم أيضاً من اختصاصاتٍ متنوعة. ويضم فريق البحث أعضاء من الأوساط الأكاديمية ومن القطاع العملي على حد سواء، مع مزيج من علماء الحاسوب والصحفيين وعلماء الاجتماع (بما في ذلك أولئك الذين لهم خلفية في دراسات الإعلام والعلوم السياسية)، واختصاصيين في حقوق الإنسان مع تركيز على حُريّة التّعبير. وهذا يؤدي إلى اعتماد مقاربة تعتمد على طائفة من المنظورات، وترتبط ارتباطاً وثيقاً بالممارسة والأثر على حد سواء. وبالتالي، هناك تركيز أيضاً على الاستجابات التقنية مثل: استخدام الذكاء الاصطناعي، بالإضافة إلى الاستجابات التثقيفية، والصحفية، والاستجابات من قطاع التكنولوجيا الصناعية.

ويُعدّ هذا التقرير سابقةً أيضاً بفعل تركيزه بشكل خاص على الاستجابات للمعلومات المُضلّلة. فكما ذُكر أعلاه، تركز تقارير بارزة أخرى على المعضلات (Witness Media Lab, 2019)، والتداعيات على السياسات (e.g. LSE, 2018; Annenberg Public Policy Center, 2020)، والآثار السياسية (Marsden & Meyer, 2019; Pamment et al., 2018)، والأهمية بالنسبة إلى الصحافة والاستجابات من مجال الصحافة أيضاً (Ireton & Posetti, 2018; Posetti et al., 2019a). كذلك، يعرض هذا التقرير مجموعة الاستجابات لكافة أنواع المعلومات المُضلّلة، بدلاً من التركيز على نوع معين مثل حملات التّضليل السياسي (Brooking et al., 2020) أو المسائل المتعلقة بالوصول إلى بيانات الشركات وكيف يؤثر ذلك على البحث الأكاديمي (Walker et al., 2019).

وهناك زاوية جديدة أخرى في هذه الدراسة، هي أن أزمة المعلومات المُضلّلة تُعالج بشكل منهجي في ضوء التحديات المتعلقة بحُريّة التّعبير، مع ما يترتب على ذلك من آثار على حُريّة الصحافة، كما هو الحال في الاستجابات التشريعية، وغيرها.

كذلك، فإنّ تصنيف الاستجابات التي طوّرتها هذه الدراسة يقسم أزمة المعلومات المُضلّلة بطريقة جديدة. فيدرس كل استجابة من مختلف الزوايا، وينظر إلى ما وراء «ما هو وكيف هو» ليعالج مسائل على غِرار «مَن يقوم بتمويل هذه الاستجابات (والآثار المترتبة عليها)؟» وما هي نقاط القوة والضعف فيها» و«ما هي نظرية التغيير التي تستند إليها؟» وتلقي هذه المقاربة ضوءً إضافياً على الافتراضات التي تستند إليها الاستجابات وعلى مدى إدماجها لعناصر الرصد والتقييم في أنشطتها.

4

الفصل الرابع:
الاستجابات الهادفة
إلى تحديد
المعلومات المُضلّلة



1.4 استجابات الرصد والتحقُّق من الوقائع

المؤلفون: دينيس تيسو، جولي بوزيتي، سام غريغوري

يستعرض هذا الفصل استجابات التحقُّق من الوقائع ويضعها في إطار نظامي كجزء من رصد المعلومات المُضلّلة (إلى جانب الاستجابات الاستقصائية على النحو الوارد في الفصل 4.2). وهنا، ينصب التركيز على استجابات التحقُّق من الوقائع ذات النطاق العالمي والإقليمي والوطني في عدد كبير من البلدان واللُّغات، والتي يمكن أن تكون إمّا عمليات مستقلة أو تابعة لوكالات إخبارية. ويرد في هذا الفصل أيضاً وصف للطريقة التي تعمل فيها هذه الجهود على التصدّي للمعلومات المُضلّلة.

تعريفات

يُعتبر التدقيق في المعلومات verification جوهر الصحافة (Kovach & Rosenstiel, 2001). والتدقيق هو تقنية تحرير يستخدمها الصحفيون والجهات المستقلة للتحقق من الوقائع، وذلك للتأكد من صحة بيان و/أو وثيقة وما شابه، وكذلك للتأكُّد من حقيقة المنصات والهُويّات (البشرية والرقمية) لمن يقوم بإنتاج المُحتوى أو تداوله. ولكن لابد من التمييز بين التدقيق والتحقُّق من الوقائع (Silverman et al., 2014):

- إن التدقيق هو مجال يصب في صميم الصحافة يتزايد استخدامه وتطبيقه في مهن أخرى.
- أما التحقُّق من الوقائع فهو تطبيق مُحدد للتدقيق – سواءً في الصحافة [أو من قبل منظمات أخرى، بما في ذلك المنظمات غير الحكومية]. وفي هذا الصدد، يعتبر التدقيق ممارسةً أساسيةً تسمح بالتحقُّق من الوقائع.

وينطوي التحقُّق من الوقائع بصورة متزايدة أيضاً على كشف الزيف بشكل استباقي عبر نشر معلومات تكشف هذا الزيف، وغالباً عبر تحديد العملية المنهجية التي استخدمت للتوصُّل إلى هذه الخلاصة.

1.1.4 ما هي الأمور ومن هي الجهات التي تستهدفها هذه الاستجابات؟

تقضي استجابات التحقُّق من الوقائع بالتدقيق ليس فقط في العملية الصحفية (ومُخرجاتها)، ولكن أيضاً في الادعاءات والتصريحات والبيانات الصادرة عن أطراف ثالثة والمتداولة خارج مجال وسائل الإعلام التقليدية، وخاصةً على شبكات التواصل الاجتماعي.

ويمثل التدقيق في صحة جهة فاعلة أو مؤسسة أو حساب على وسائل التواصل الاجتماعي النقطة التي تبدأ منها عملية التحقُّق من الوقائع في تغذية الاستجابات الاستقصائية (مراجعة القسم 4.2). وتشكّل الاستجابات المتعلقة بتحديد المعلومات المُضلّلة، على غرار استجابات الرصد والتحقُّق من الوقائع، أساساً للتحقيق في منشأ المعلومات المُضلّلة وانتشارها، ممّا يساهم في بناء قاعدة أدلة تعتمد عليها الأنواع الأخرى من الاستجابات. وسترد أمثلة محددة في القسم 4.1.4 أدناه.

2.1.4 مَن هي الجهات التي تسعى إلى مُساعدتها استجابات الرصد والتحقّق من الوقائع؟

إنّ التحقُّق من الوقائع يسمح لشركات التواصل عبر الإنترنت[40] بكشف المعلومات المُضلّلة وتطوير الاستجابات التي تقلّص أو تزيل من ظهورها و/أو مصداقيتها. كذلك، يُساعد التدقيق أيضاً الحكومات والمنظمات الدولية على تحديد نوعية الإجراءات الواجب اتخاذها ومتى يجب اتخاذها وإذا كان هناك حاجة لذلك. على سبيل المثال، إطلاق سياسات أو مبادرات عملية مثل الحملات الهادفة للتصدي للمعلومات المُضلّلة. وأخيراً، تزوّد عمليات التحقُّق من الوقائع التي تُنشَر نتائجها؛ المواطنين، بمصادر مفيدة للمعلومات الموثوقة.

3.1.4 ما هي النتائج التي يتم نشرها؟

تنشر هذه الاستجابة استنتاجاتها أي ما هي الأمور التي تم التحقُّق منها، وكيف، وما هو تقييمها لفائدة الصحة أو الزيف، وهل هي غير قابلة للتحقق أو غيره (مثلاً الآراء التي لا يمكن التحقُّق منها في حد ذاتها، على الرغم من أنه متى كانت مبنيةً على وقائع مزعومة، تكون بعض جوانبها قابلةً للتحقق في ما يخص مدى زيف هذه «الوقائع» أو تضليلها). ومن المسلّم به أنّ مستخدمي وسائل التواصل الاجتماعي يميلون إلى مشاركة مُحتوى عمليات التحقُّق من الوقائع المنشورة أقل بكثير من مشاركتهم للمعلومات المُضلّلة واسعة الانتشار التي يتم كشف زيفها من خلال عمليات التحقُّق (Shin & Thorsen, 2017). كما وتبرز مخاوف من أن يؤدي لفت الانتباه إلى الأكاذيب إلى تضخيمها. ومع ذلك، فإن الافتراض السائد هو أن عملية التحقُّق وكشف الزيف لا تزال وسيلةً أساسيةً لكشف الحقيقة ومحاسبة الأفراد والشخصيات العامة والمؤسسات ووسائل الإعلام لنشرها ادعاءات غير صحيحة (Sippitt, 2020; Friedman, 2020; Qui, 2020).

4.1.4 من هي الجهات الفاعلة الرئيسية ومن يمول هذه الاستجابات؟

أ. الاستجابات العالمية

منظمة First Draft

من أُولى المبادرات العالمية التي تُركّز على التحقُّق من مُحتوى وسائل التواصل الاجتماعي على المستوى الدولي التحالف غير الربحي المسمى First Draft، والمسجل في المملكة المتحدة منذ حزيران/يونيو 2015. وكان الهدف من تحالف First Draft عند إنشائه تقديم إرشادات عملية وأخلاقية لوسائل الإعلام حول تحديد محتويات الشبكات الاجتماعية والتحقُّق منها ونشرها، وخاصةً في سياق الأخبار العاجلة.

في أيلول/سبتمبر 2016، تم توسيع النواة الأصلية التي كانت تضم تسعة شركاء (BellingCat, Dig Deeper, Emergent.info, EyeWitness Media Hub, Google News Initiative, Meedan, Reported.ly, Storyful, and VerificationJunkie) ليصبح التحالف شبكةً دوليةً تضم منظمات إعلامية وأكاديميين ومنصات تواصل اجتماعي ومنظمات مجتمع مدني. وفي الوقت نفسه، انضمت منظمة First Draft إلى مشروع ProPublica من ElectionLand الذي يهدف إلى تحديد وتتبّع تعرّض الناخبين إلى المعلومات المغلوطة والمُضلّلة.

[40] https://www.disinfo.eu/resources/covid-19/platforms-responses-to-covid-19-mis-anddisinformation

خـلال الانتخابـات الرئاسـية الأميركيـة 2016. وقـد تميّـز العمـل علـى هـذا المشـروع بالتعـاون مـع طـلاب مـن 13 مدرسـة صحافيـة تـم تدريبهـم علـى تقنيـات جمـع الأخبـار الاجتماعيـة وتقنيـات التدقيـق. وقـد حظـي مشـروع Electionland بدعـم مالـي مـن قبـل مختبـر أخبـار غوغـل والمتبـرع الأمريكـي «كريـغ نيومـارك».

بعد ذلك، أطلقت منظمة First Draft عدة برامج تعاونية لرصد الانتخابات في فرنسا والمملكة المتحدة وألمانيا والبرازيل ونيجيريا. ويقوم CrossCheck، وهو تحالف لوسائل الإعلام الإخبارية ناتج عن هذا المشروع؛ يُعنى بالتحقُّق من الوقائع، برصد الشائعات التي تنتشر داخل هذه البلدان، وبنشر تقارير تكشف زيف المعلومات الكاذبة، وذلك بهدف تزويد الناخبين بوسائل تسمح لهم بالتوصُّل إلى استنتاجات بأنفسهم بدلاً من الانجرار وراء المعلومات المُضلّلة. (لمعرفة المزيد عن الاستجابات الانتخابية، يُرجى مراجعة القسم 5.3).

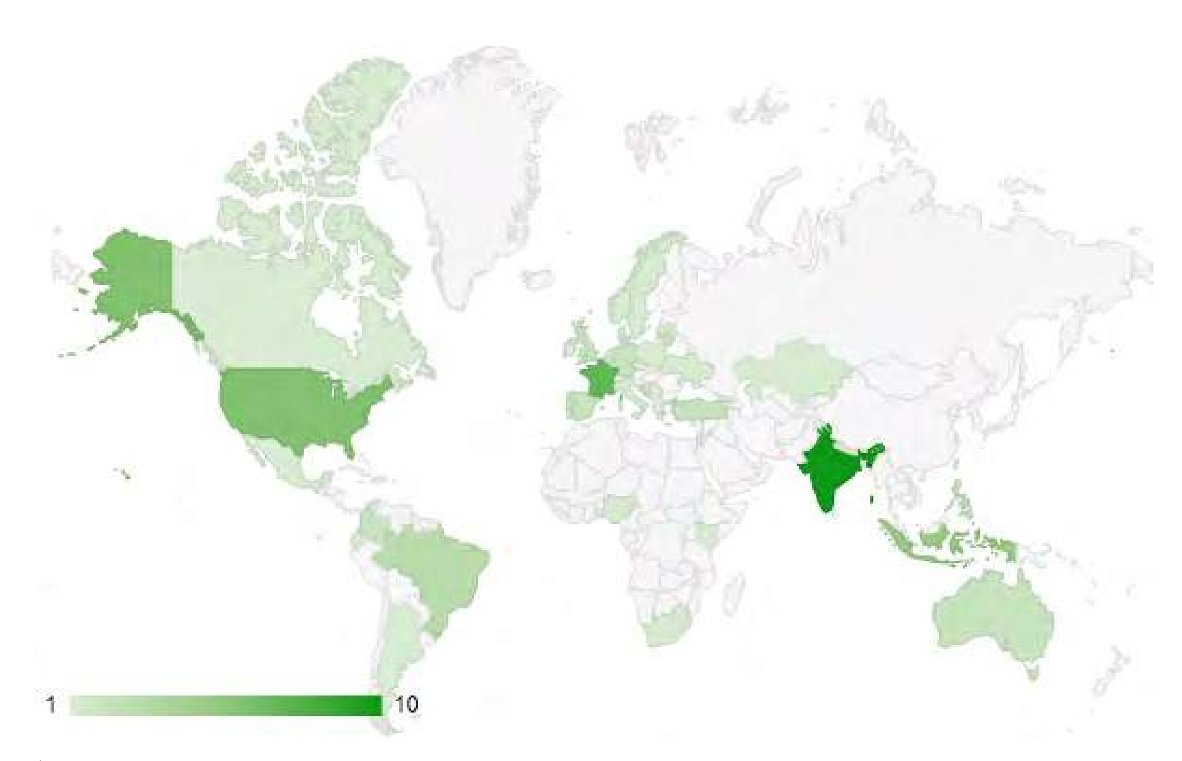
وفـي عـام 2020، سـعت منظّمـة First Draft إلـى توسـيع عملياتهـا لتطـال الأرجنتين وأسـتراليا وكندا وإندونيسـيا وجنـوب أفريقيـا وإسـبانيا وأوروغـواي، كمـا سـعت إلـى تنسـيق مشـروع عابـر للحـدود للتحقيـق فـي تكتيـكات واتجاهـات التّضليـل فـي أوروبـا خـلال فتـرات هـي أبعـد مـن الفتـرات الانتخابيـة.

وقـد حصلـت First Draft علـى منـح وتبرّعـات مـن العديـد مـن المؤسسـات الخيريـة مـن جهـات غيـر الشـريك المؤسـس فيهـا وهـي مبـادرة أخبـار غوغـل Google News Initiative ومشـروع فيـس بـوك للصحافـة Journalism Project وتويتـر. وبعـد انضمامهـا بفتـرة وجيـزة إلـى «مركـز شـورنشـتاين للإعـلام والسياسـة والسياسـة العامـة» Shorenstein Center for Media, Politics and Public Policy فـي «مدرسـة كينيـدي» بـ«جامعـة هارفـارد» فـي تشـرين الأول/أكتوبـر 2017، أصبحـت منظمـة First Draft تعمـل الآن بشـكل مسـتقل، وتعتمـد فـي المقـام الأول علـى التمويـل مـن شـركات التواصـل عبـر الإنترنـت. وتـرد تفاصيـل عـن المبـادرات التعاونيـة المتعلقـة بالانتخابـات التـي نظمتهـا First Draft فـي القسـم 5.3.

الشبكة الدولية للتحقق من الوقائع (IFCN)

تـم إطـلاق الشـبكة الدوليـة للتحقـق مـن الوقائـع (IFCN, 2019a) فـي أيلول/سـبتمبر 2015 كوحـدة أعمـال داخـل كليـة الصحافـة غيـر الربحيـة فـي «معهـد بوينتـر للدراسـات الإعلاميـة»، ومقـره فـي سـانت بيترسـبرغ فـي فلوريـدا، الولايـات المتحـدة الأميركيـة. وقـد أطلـق المعهـد، الـذي يملـك «صحيفـة تامبـا بـاي تايمـز»، هـذه الشـبكة لجمـع الجهـات التـي تعمـل فـي التحقُّـق مـن الوقائـع مـن جميـع أنحـاء العالـم وتعزيـز الممارسـات الجيـدة وتبـادل المعـارف فـي هـذا المجـال.

وتتمثل مهمة الشبكة في رصد الاتجاهات والتنسيقات وصنع السياسات المتعلقة بالتحقُّق من الوقائع في جميع أنحاء العالم، ونشر مقالات منتظمة عن التحقُّق من الوقائع، وتعزيز التدريب – سواءً أكان ذلك حضورياً أو عبر الإنترنت – فضلاً عن ضمان المعايير الأساسية من خلال مدونة مبادئ لجهات التحقُّق من الوقائع. وفي 6 آب/أغسطس 2020، كان للشبكة 79 جهة تم التحقُّق منها؛ تعد من الموقّعين الفعّالين على مدونة المبادئ و14 جهةً تم التحقُّق منها وتوقيعها قيد التجديد (IFCN, 2020d). وتوضّح الخريطة أدناه التوزيع الجغرافي للموقعين، وبعضهم من جهات التحقُّق من الوقائع على المستويين الوطني والدولي على حد سواء. وعمليـة التأكـد مـن جهـات التحقُّـق مـن الوقائـع مهمـة، لأنـه مـن الممكـن اليـوم، فـي هـذه السـاحة المتنـازع عليهـا، أن تقـوم مبـادرات ناقصـة أو حتـى زائفـة باسـتغلال هـذه التسـمية لأغـراض بعيـدة كل البعـد عـن التصـدي للأكاذيـب.



الشكل 4. خريطـة جغرافيـة للموقعيـن علـى مدونـة مبـادئ الشـبكة الدوليـة للتحقُّـق مـن الوقائـع (67 موقّعـاً فعّـالاً و14 قيـد التجديـد فـي أوائـل عـام 2020)

وعلى الموقعين التقيّد بخمسة التزامات (IFCN, 2019c):

1. الالتزام بعدم التحزب والإنصاف؛
2. شفافية المصادر؛
3. شفافية التمويل والتنظيم؛
4. شفافية المنهجية؛
5. سياسة تصحيح مفتوحة ونزيهة.

وقد تمّ إطلاق مدونة المبادئ هذه في أيلول/سبتمبر 2016، بعد عام واحد من نشأة الشبكة. وفي العام 2017، اعتمدت الشبكة عمليةً خاصة بتقديم الطلبات وفحص المُنتسبين وذلك بعد إعلان فيس بوك أنّ الشرط الأدنى لقبول جهة ما كطرف ثالث للتحقق في الوقائع أن تكون هذه الجهة قد وقّعت على هذه المدونة[41].

وتلعب الشفافية، التي غالباً ما تُعتبر في أدبيات الدراسات الإعلامية مبدئاً أخلاقياً جديداً في الصحافة، دوراً مهماً في هذه الالتزامات. وهذا يتقاطع مع ظهور الشفافية بين المدونين وبين أولى جهات التحقُّق من الوقائع كبديل ضروري أو طبيعي للمثال الصحفي الأعلى، وهو الموضوعية (Graves, 2013). وهذا مبني على فكرة اعتنقها الفيلسوف «ديفيد واينبرغر» عام 2009 حيث أكد أنّ «الشفافية هي الموضوعية الجديدة» (Weinberger, 2009). وقد ترسّخت الآن فكرة الشفافية وصلتها بالثقة في الصحافة الموثوقة كمعيار في عمليات التحقُّق من الوقائع كما وفي الصحافة المحترفة. ويمكن للشفافية التي تتيحها التفسيرات المنشورة لعمليات التدقيق والتحقُّق من الوقائع؛ أن تُعزّز من قدرة أي عمل على التصدي لادعاءات التحيُّز أو عدم الصحة، حيث أنه تم الإفصاح عن قاعدة الإثبات التي اتبعتها في عملية التحقُّق من الوقائع.

[41] لمزيد من التفصيل، يُرجَى مُراجعة القسمين 4.2 و7.1.

وتنظم الشبكة مؤتمراً دولياً سنوياً (Global Fact) لتعزيز الجهود التعاونية بين كافة جهات التحقُّق من الوقائع في العالم. وفي نسخة 2019، عُقد مؤتمر Global Fact 6، في كيب تاون، جنوب أفريقيا، بحضور أكثر من 250 مشاركاً يمثلون 55 بلداً و146 منظمة فاعلة. أما Global Fact 7، الذي كان من المقرر انعقاده في أوسلو، النرويج، في حزيران/يونيو 2020، فقد عُقد في نهاية المطاف عبر الإنترنت بسبب جائحة فيروس كورونا. وتموّل الشبكة أيضاً زمالات سنوية وصندوق Fact Forward Fund وهو مُبادرة للابتكار في مجال التحقُّق من الوقائع وبرنامج للتمويل الجماعي crowdfunding. وأخيراً، تدعو الشبكة إلى توسيع نطاق جهود التحقُّق من الوقائع على الصعيد العالمي، بما في ذلك من خلال تخصيص يوم عالمي للتحقُّق من الوقائع، هو 2 نيسان/أبريل من كل عام.

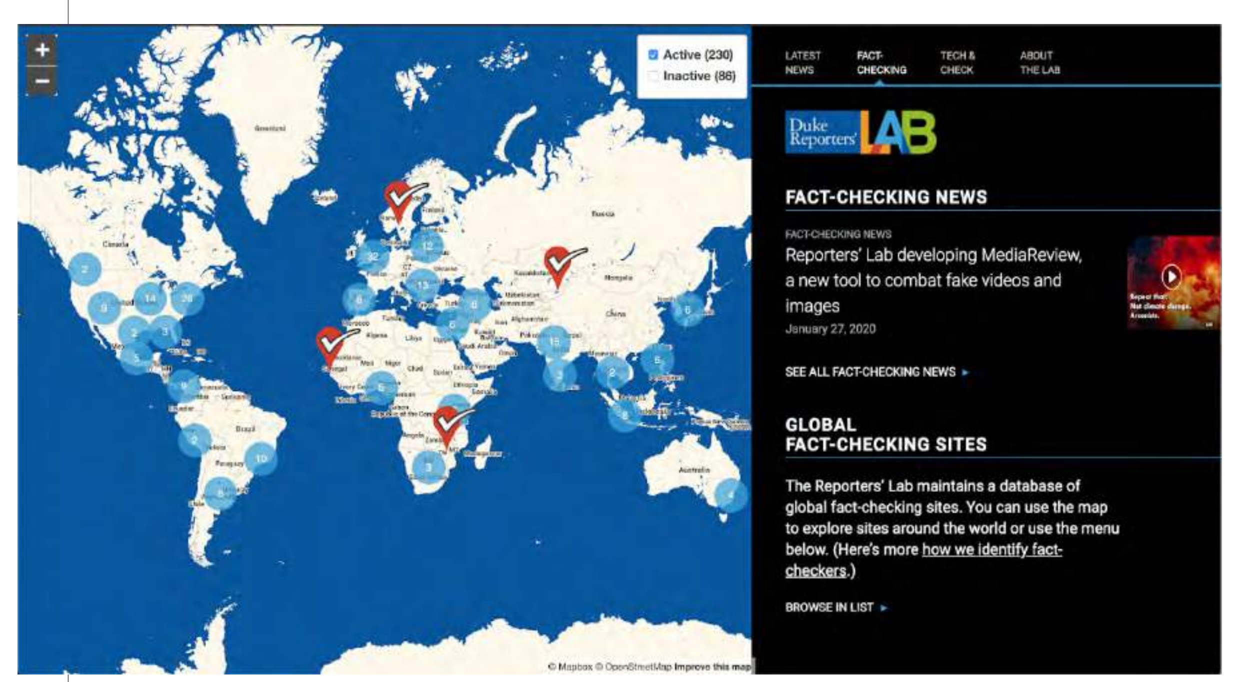
وتلقّت الشبكة تمويلاً من مؤسسة Arthur M. Blank Family Foundation، ومن مختبر «جامعة ديوك» للمراسلين Duke Reporters' Lab، ومن مؤسسة Melinda Gates Foundation، ومن غوغل، ومن National Endowment for Democracy ومن شبكة Omidyar، ومن شبكة Open Society Foundations، ومن مؤسسة Park Foundation كما ورد في (IFCN, 2019d).

قاعدة بيانات مختبر «جامعة ديوك» للمُراسلين

إنّ مختبر المراسلين هو مركز لأبحاث الصحافة في «كلية سانفورد للسياسة العامة» بجامعة ديوك في الولايات المتحدة. وكان من أحد مشاريعه الرئيسية إنشاء قاعدة بيانات عالمية[42] عن العمليات الرئيسية للتحقق من الوقائع، الفعّالة أو غير الفعّالة، وبالتالي توثيق صعود قطاع التحقُّق من الوقائع، في كل بلد.

وبغض النظر عن الرسم الجغرافي الذي يعرض جميع المنظمات العاملة في التحقُّق من الوقائع، تسمح قاعدة البيانات للمُستخدم بتصفُّح المُحتوى حسب القارات والبلدان، كما ويتم تحديث المُحتوى بصورة دوريّة. ولإضافة مواقع جديدة للتحقق من الوقائع إلى قاعدة البيانات لابد لهذه المواقع أن تستوفي معايير عدم التحزب وأن تتحقق من الوفاء بالوعود السياسية (مثلاً، البيانات الحزبية أثناء الانتخابات)، وأن تعتمد الشفافية بشأن المصادر والأساليب، وكذلك بشأن التمويل والانتماء، وأن تكون رسالتها الرئيسية هي الأخبار والمعلومات. وفي نيسان/أبريل 2020، كانت قاعدة بيانات مختبر المراسلين تضم 237 موقعاً فعّالاً و91 موقعاً غير فعّال حول العالم في 78 دولة.

42 https://reporterslab.org/fact-checking/



الشكل **5.** صورة لقاعدة بيانات التحقُّق من الحقائق التابعة لمُختبر المُراسلين في «جامعة ديوك»

شبكة التحقُّق من الوقائع من قبل أطراف ثالثة التابعة لفيس بوك

تستخدم شركات التواصل عبر الإنترنت عادةً عمليات داخلية أو خارجية للتحقق من الوقائع، تسمح بتصميم استجاباتها الخاصة بتنظيم المُحتوى، على سبيل المثال، لإزالة منشورات معيّنة أو تخفيض أهميتها أو إخفائها. ويرد وصف لهذه الأمور بمزيد من التفصيل في القسم 6.1 من الفصل 6.
ومن بين أنظمة الشركات، يستحق نظام فيس بوك الاهتمام باعتباره البرنامج الدولي الوحيد واسع النطاق «للتدقيق من قبل أطراف ثالثة» بين شركات التواصل عبر الإنترنت، وقد تم إطلاقه بعد فترة وجيزة من الانتخابات الرئاسية الأميركية عام 2016 (Zuckerberg, 2016a). وقد أعلن الرئيس التنفيذي لفيس بوك «مارك زوكربيرغ» عن المشروع على صفحته الخاصة على فيس بوك، فأكد على أنّ فيس بوك يأخذ مسألة «المعلومات المُضلّلة على محمل الجد»، مقراً بوجود العديد من منظمات التحقُّق من الوقائع المحترمة التي تسعى شركته للتعلّم منها. وكان قد ذكر في السابق، أنّ أكثر من 99% مما يراه الناس على فيس بوك هو مُحتوى أصيل وصحيح (Zuckerberg, 2016b). وتم تفسير إعلانه عن مبادرة التحقُّق من الوقائع من قبل أطراف ثالثة على أنه محاولة لمواجهة الانتقادات التي وُجهت إلى فيس بوك بعدم تدخل الشركة لوقف انتشار المعلومات المُضلّلة خلال الحملة الانتخابية الرئاسية الأميركية عام 2016.

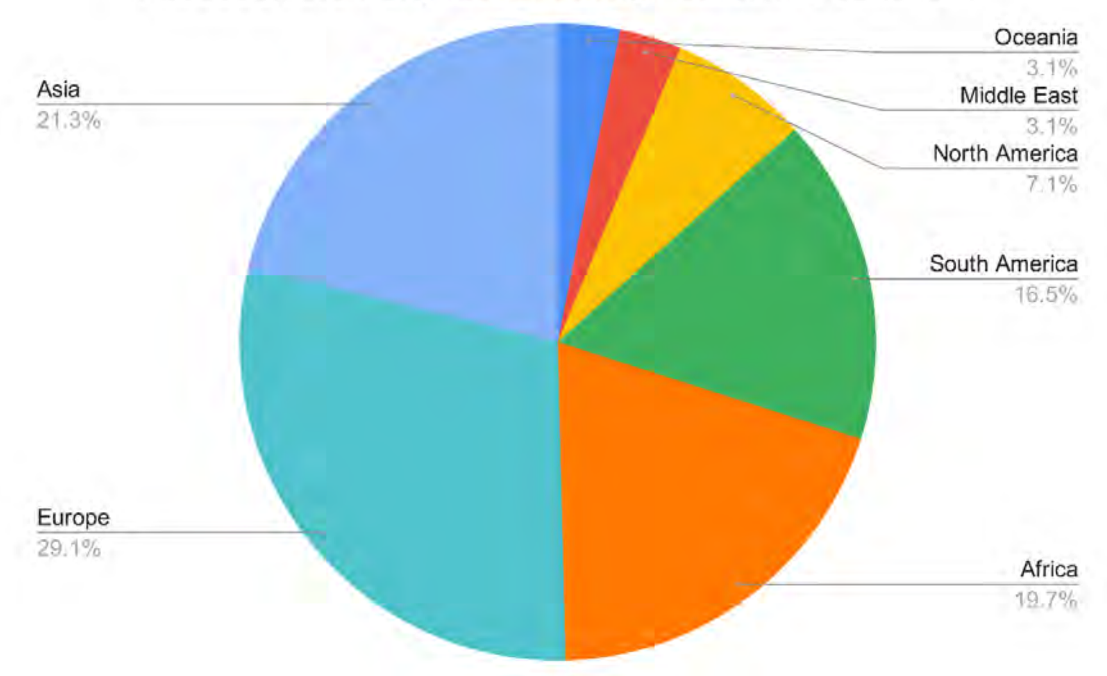
وبعد شهر واحد من الانتخابات الأميركية عام 2016، أعلن فيس بوك عن إطلاق برنامج يهدف إلى العمل مع أطراف ثالثة تعمل في مجال التحقُّق من الوقائع؛ كانت قد وقعت على مدونة مبادئ الشبكة الدولية للتحقق من الوقائع (IFCN) التابعة لمعهد بوينتر (Mosseri, 2016). ويقوم برنامج فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة بالتعاقد مع المنظمات لمراجعة وتقييم صحة المُحتوى، بما في ذلك القصص والإعلانات غير السياسية (مراجعة المناقشة أدناه حول القيود على التحقُّق من المُحتوى السياسي في هذه البرنامج).

وبمجرد أن يقوم هؤلاء الشركاء العاملين في مجال التحقُّق من الوقائع بتصنيف قصة ما على أنها «خاطئة»، يضعها فيس بوك في مرتبة أدنى على شريط الأحداث Newsfeed ما لم تتم مُراجعتها من قبل فيس بوك

بناءً على السياسات والإجراءات و/أو المدفوعات المرتبطة بالعقود التي تعمل بموجبها منظمات التحقُّق من الوقائع (Pasternack, 2020). (عموماً، يتم استبعاد المُحتوى المبني على الرأي، وفئات معينة من الإعلانات السياسية والخطاب السياسي من قبل السياسيين والأحزاب السياسية والجهات ذات العلاقة بها). وعلى انستغرام، يصعّب فيس بوك العثور على المُحتوى الذي تم التحذير منه بناءً على ووفقاً لهذا البرنامج من خلال انتقائيته من صفحات الاستكشاف والوسم «هاشتاج»، وتخفيض مُعدّل ظهوره على شريط الأحداث الجارية. كذلك تتم الإشارة بشكل بارز إلى المُحتوى المصنف على أنه خاطئ كلياً أو جزئياً على فيس بوك وانستغرام[43] حتى يتمكن الناس من اتخاذ قرار أفضل في ما يخص ما يودون قراءته والوثوق به ومشاركته. ويتم وضع التنبيه أعلى الصور ومقاطع الفيديو غير الصحيحة كلياً أو جزئياً، بما في ذلك مُحتوى «القصص» على انستغرام، ويتم ربطها بالتقييم الذي وضعته الجهة التي قامت بالتحقُّق من الوقائع[44].

وقبل كانون الأول/ديسمبر 2017، في حال قيام منظمة تحقُّق من الوقائع بتصنيف قصة ما على أنها خاطئة (أو «زائفة» وفقًا لبروتوكول فيس بوك)، كانت هذه المنظمة تقوم بتبليغ فيس بوك، فيتم وضع شارة تحذير تصفها «بغير المتفق عليها»، ويتم ربطها بالمقالة المقابلة لها (على موقع جهة التحقُّق من الوقائع) لتوضيح سبب التصنيف. ووفقاً لفيس بوك، من شأن هذا أن يحد من ظهور هذه المنشورات بنسبة 80% (Lyons, 2018a) وبالتالي يُساعد على احتواء انتشارها. ومع ذلك، تستغرق هذه العملية حوالي ثلاثة أيام بعد نشر المُحتوى للمرة الأولى (Silverman, 2017b). ويؤكّد فيس بوك على أنه يستخدم أيضاً المعلومات الصادرة عن جهات التحقُّق من الوقائع لتحسين تقنيته الهادفة إلى الإمساك بالمُحتوى الكاذب بوتيرةٍ أسرع. ويمكن الاطلاع على مزيد من التقييم لعملية تصنيف المُحتوى في الفصل 7.3.





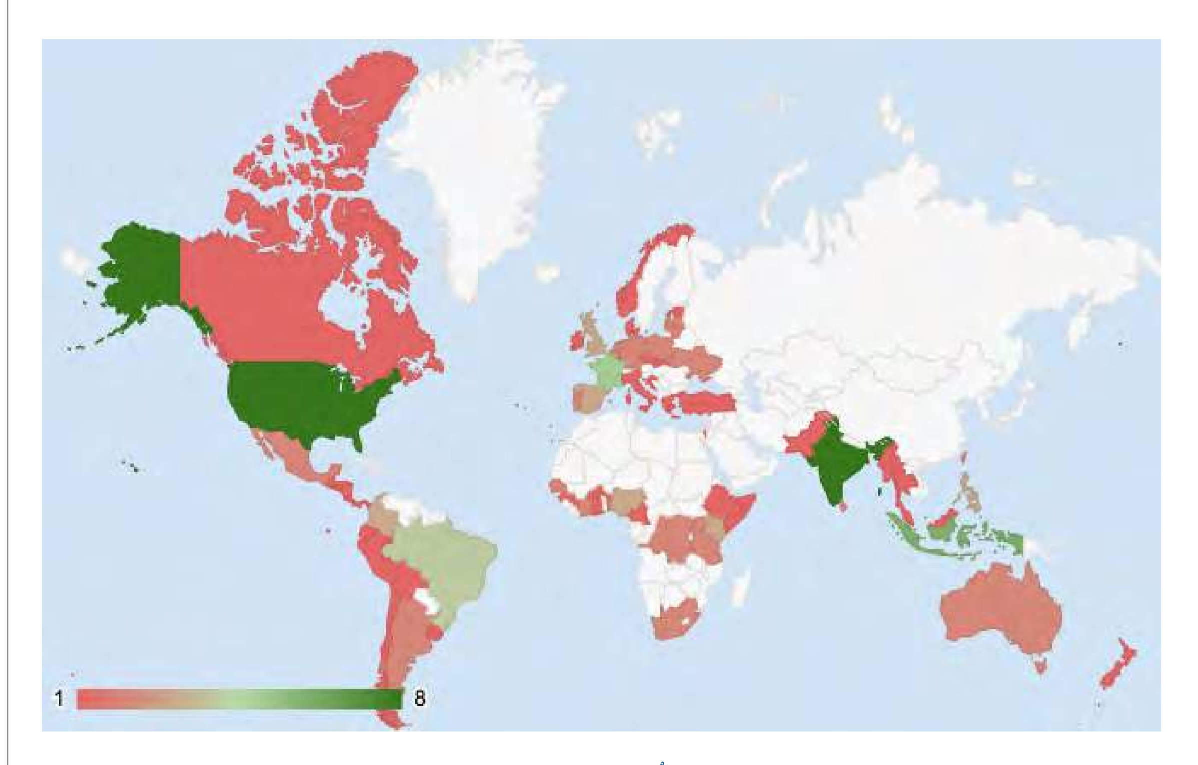
الشكل 6. صورة لشبكة التحقُّق من الوقائع من قبل أطراف ثالثة التابعة لفيس بوك مُقسَّمة حسب القارات والعمليات

43 https://about.fb.com/wp-content/uploads/2020/06/Elections-Fact-Sheet.pdf

44 https://about.fb.com/news/2019/10/update-on-election-integrity-efforts/

لقـد تـم إطـلاق برنامـج التحقُّـق مـن الوقائـع مـن قبـل أطـراف ثالثـة فـي الولايـات المتحـدة فـي منتصـف كانـون الأول/ديسـمبر 2016 بالتعـاون مـع وكالـة The Associated Press، وPolitifact، وFactCheck.org، وSnopes، ومحطـة ABC News. ثـم توسّـع بسـرعة على صعيـد دولـي. فـي حزيران/يونيـو 2018، وبعـد ثلاثـة أشـهر مـن فضيحـة Cambridge Analytica، قـام البرنامـج بربـط 25 منظمـة فـي 14 دولـة.

وفـي أوائـل أيلول/سـبتمبر 2020، عندمـا تـم الانتهـاء مـن هـذا التقريـر البحثـي، عقـد فيـس بـوك شـراكةً مـع 74 منظمـةً خارجيـةً للتحقـق مـن الوقائـع فـي جميـع أنحـاء العالـم، بأكثـر مـن 50 لغـة (يسـتند تحليـل البيانـات هـذا إلـى قائمـة شـركاء فيـس بـوك السـابقة وأحـدث خريطـة لشـركائه[45]).



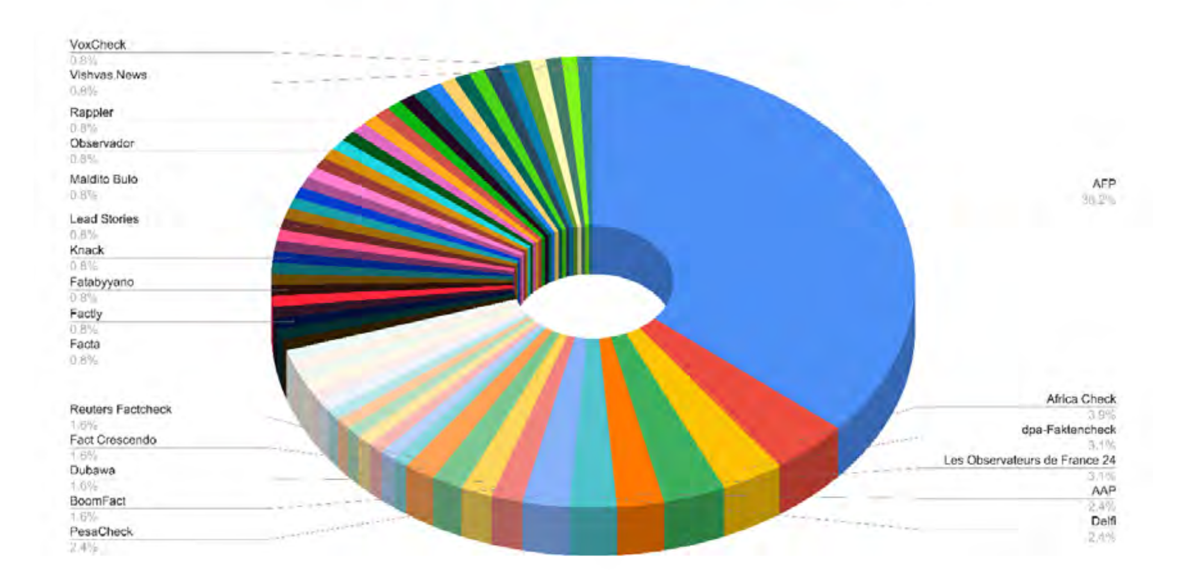
الشكل 7. خريطة للتوزيع العالمي لشبكة التحقُّق من الوقائع من قبل أطراف ثالثة التابعة لفيس بوك

توضـح الخريطـة أعـلاه (الشـكل 7) حالـة برنامـج التحقُّـق مـن الوقائـع مـن قبـل أطـراف ثالثـة التابـع لفيـس بـوك فـي أيلول/سـبتمبر 2020. ويبيـن الجـدول 2 أدنـاه التوزيـع حسـب عـدد عمليـات التحقُّـق مـن الوقائـع التـي تعاقـد عليهـا فيـس بـوك فـي 10 أيلول/سـبتمبر 2020.

45 https://web.archive.org/web/20200728165712if_/https://www.facebook.com/business/help/182222309230722
تم إيقافها من قِبل فيس بوك في آب/أغسطس 2020 واستُعيض عنها بخريطة:
https://www.facebook.com/journalismproject/programs/third-party-fact-checking/partner-map

التغطية	عدد العمليات
الولايات المتحدة الأميركية	10
الهند	8
إندونيسيا	6
فرنسا، إسبانيا	5
البرازيل	4
ألمانيا، بلجيكا، جمهورية الكونغو الديمقراطية، كولومبيا، كينيا، ميانمار، نيجيريا، الفلبين، المملكة المتحدة	3
الجزائر، الأرجنتين، أستراليا، النمسا، البحرين، كندا، الجمهورية التشيكية، مصر، إثيوبيا، العراق، ساحل العاج، الأردن، الكويت، لاتفيا، لبنان، ليبيا، ليتوانيا، المكسيك، المغرب، هولندا، نيوزيلاندا، سلطنة عمان، فلسطين، بيرو، بولندا، البرتغال، قطر، المملكة العربية السعودية، السنغال، سنغافورة، جنوب أفريقيا، سيريلانكا، السودان، سويسرا، سوريا، جمهورية تنزانيا المتحدة، تونس، تركيا، أوغندا، أوكرانيا، الإمارات العربية المتحدة، اليمن	2
أذربيجان، بنغلاديش، بنين، بوليفيا، بوركينا فاسو، بوروندي، الكاميرون، جمهورية أفريقيا الوسطى، تشيلي، جزر كوك، كوستاريكا، كرواتيا، الدنمارك، إكوادور، السلفادور، إستونيا، فيجي، بولينيزيا الفرنسية، غانا، اليونان، غواتيمالا، غينيا كوناكري، هندوراس، آيرلندا، إسرائيل، إيطاليا، كيريباتي، لوكسمبورغ، مقدونيا، ماليزيا، مالي، جزر مارشال، ميكرونيزيا، الجبل الأسود، ناورو، كاليدونيا الجديدة، نيكاراغوا، نيوي، النرويج، باكستان، بالاو، بنما، بابوا غينيا الجديدة، باراغواي، ساموا، سلوفاكيا، جزر سُليمان، الصومال، جمهورية كوريا، السويد، تايلاند، تونغا، توفالو، أوروغواي، فانواتو، فنزويلا وزامبيا	1

الجدول 2. توزيع شبكة التحقُّق من الوقائع من قبل أطراف ثالثة التابعة لفيس بوك وفقاً للبلد وعدد العمليات



الشكل 8. توزيع برنامج التحقُّق من الوقائع من قبل أطراف ثالثة التابع لفيس بوك وفقاً للمنظمات المعنية

ويتمّ اختيـار جهـات التحقُّـق من الوقائـع والدفـع لهـا[46] من قبل فيـس بـوك. وكشـرط أسـاسـي، لابـد أن تكـون كافـة هـذه الجهـات مصـادق عليهـا مـن قبـل الشـبكة الدوليـة للتحقـق مـن الوقائـع وأن تلتـزم بمدونـة المبـادئ الخاصـة بهـا. ويمثّل هـذا البرنامـج أكبـر شـبكة دوليـة منظمـة للتصـدي للمعلومـات المُضلّلـة ويغطـي كل مـن فيـس بـوك وانسـتغرام (منـذ أيار/مايـو 2019)[47]، ولكنـه ذو فائـدة علـى نطـاق أوسـع، حيـث أن المُحتـوى الـكاذب أو المضلـل علـى فيـس بـوك غالبـاً مـا يتـم نشـره عبـر الشـبكات الاجتماعيـة الأخـرى مثـل تويتـر، يوتيـوب أو واتسـاب المملـوك مـن قِبـل فيـس بـوك. لـذا، فـإن الحـد مـن انتشـار المعلومـات المُضلّلـة علـى فيـس بـوك وانسـتغرام يمكـن أن يحـد نظريـاً مـن «قفـز» هـذه المعلومـات إلـى واتسـاب ومـا بعـده. وعلـى الرغـم مـن أن واتسـاب لا يُرسـل مباشـرة المحتويـات غيـر المتفـق علـى صحتهـا إلـى جهـات التحقُّـق مـن الوقائـع، إلا إنـه يحتـوي علـى روبـوت دردشـة chatbot (غيـر معلـن عنـه إلا قليـلاً) يسـمح للمسـتخدمين بالحصـول علـى نصائـح حـول التحقُّـق مـن الوقائـع ويربطهـم بجهـات التحقُّـق المحليـة عبـر قـاعـدة بيانـات الشـبكـة الدوليـة للتحقـق مـن الوقائـع ليقومـوا، إن رغبـوا، بالاتصـال بهـا للتحقـق مـن موضـوع معيّـن[48].

واسـتجابةً لجائحـة فيـروس كورونـا، قـادت الشـبكة أيضـاً عمليـة تطويـر روبـوت دردشـة علـى واتسـاب يتيـح للمسـتخدمين البحـث فـي قاعـدة بياناتهـا المخصصـة للتحقـق مـن الوقائـع حـول الفيـروس (Grau, 2020). بالإضافـة إلـى ذلـك، وكمـا سـيُذكر بمزيـد مـن التفصيـل فـي الفصـل 7.3، بـدأ واتسـاب فـي آب/أغسـطس 2020 باختبـار ميـزة جديـدة (فـي سـتة بلـدان) تتيـح للمسـتخدمين القيـام بأنفسـهم بتحقـق بسـيط مـن وقائـع الرسـائل واسـعة الانتشـار، مـن خـلال البحـث عـن المُحتـوى علـى غوغـل (Sweney, 2020). وفـي مـا يخـص

46 يُرجَى مراجعة المناقشة أدناه حول قضايا الشفافية في ما يتعلق بالرسوم.

47 https://about.fb.com/news/2019/12/combatting-misinformation-on-instagram/

48 https://faq.whatsapp.com/general/ifc-n-fact-checking-organizations-on-whatsapp

انستغرام، تتم إزالة المنشورات التي تصنّفها جهات التحقُّق على أنها غير صحيحة من نتائج الاستكشاف وصفحات الوسم «هاشتاغ». وبالإضافة إلى ذلك، تخفّض مرتبة المحتويات المصنفة كخاطئة على شريط أحداث وقصص انستغرام[49].

وفي الفترة التي سبقت الانتخابات الرئاسية الأميركية عام 2020، أصبحت الولايات المتحدة الدولة الأولى من حيث عدد منظمات التحقُّق من الوقائع الذي بلغ عشر منظمات، تليها الهند (ثماني منظمات) التي شهدت انتشاراً للمعلومات المُضلّلة من خلال تناقل المعلومات شفهياً أو عبر واتساب؛ ممّا أدى إلى تأجيج عنف العصابات وارتفاع الخسائر البشرية (McLaughlin, 2018)، وإندونيسيا (ست منظمات).

ويُظهر توزيع الرسم البياني بناءً على المنظمات (في الشكل 8) بوضوح، أن وكالة الأنباء الفرنسية قد حصلت على حصة رائدة في البرنامج من خلال إطلاق عمليات تحقق من الوقائع في أكثر من 70 دولة، وبالعمل مع 90 صحافياً (بما في ذلك الأرجنتين وأستراليا وبلجيكا وبوليفيا والبرازيل وبوركينا فاسو والكاميرون وكندا وتشيلي وكولومبيا وكوستاريكا والجمهورية التشيكية وساحل العاج وجمهورية الكونغو الديمقراطية والإكوادور والسلفادور وفرنسا وإثيوبيا وألمانيا وغواتيمالا وهندوراس والهند وإندونيسيا وكينيا وماليزيا والمكسيك وهولندا ونيكاراغوا والعديد من بلدان الشرق الأوسط وشمال أفريقيا (بشكل رئيسي، لبنان) ونيجيريا وباكستان وبنما وبيرو والفلبين وبولندا والسنغال وسنغافورة وسلوفاكيا والصومال وجنوب أفريقيا وإسبانيا وسريلانكا وسويسرا وجمهورية تنزانيا المتحدة وتايلاند وتونس، وأوغندا والولايات المتحدة الأمريكية وأوروغواي).

وقد أوضحت وكالة الصحافة الفرنسية في كانون الأول/ديسمبر 2018 أنها جعلت التصدّي للمعلومات المُضلّلة في صلب مهمتها، وحثت على التزام وكالات الأنباء الأخرى بكشف القصص الكاذبة والمُتلاعَب بها (Fries, 2018).

وتشمل وسائل الإعلام التقليدية المعروفة الأخرى التي تتبنى التحقُّق من الوقائع وكشف الزيف، والمنضوية إلى برنامج فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة؛ المؤسسات التالية: وكالة The Associated Press (في الولايات المتحدة)، ووكالة The Australian Associated Press (في أوقيانوسيا)، رويترز، ووكالة الصحافة الألمانية (DPA؛ في ألمانيا وبلجيكا ولوكسمبورغ)، فريق المراقبين التابع المؤسسة الدولية الفرنسية France24 (في 4 بلدان: ساحل العاج وجمهورية الكونغو الديمقراطية وغينيا كوناكري وفرنسا)، وRappler (الفلبين)، The Quint (الهند)، والصحيفة الفرنسية Libération (فرنسا).

من جهة أخرى، على الرغم من أهمية الشفافية في عمليات التحقُّق من الوقائع الموضحة أعلاه، إلا أن الشفافية محدودة للغاية في ما يخص الأموال التي يدفعها فيس بوك للجهات الشريكة العاملة على التحقُّق من الوقائع. وفي تقرير نشر في تموز/يوليو 2019، أقرت Full Fact وهي منظمة بريطانية للتحقق من الوقائع بأنها تلقت 171,800 جنيه استرليني (مقابل 96 عملية تحقق من الوقائع) خلال الأشهر الستة الأولى من عملها ضمن الشراكة (Hazard Owen, 2019). وتعتمد الأموال المكتسبة جزئياً

[49] https://about.instagram.com/blog/announcements/coronavirus-keeping-people-safe-informedand-supported-on-instagram/

على مقدار التحقُّق من الوقائع الذي يتم في إطار البرنامج. وأوضحت الصحيفة الفرنسية Liberation وخدمة التحقُّق من الوقائع الخاصة بها checknews.fr أنهما حصلتا على 240000 دولار أميركي لقاء 249 مقالة تم تحميلها على فيس بوك (Checknews, 2019). أما FactCheck.org (الولايات المتحدة) فقد حصّلت مبلغ 242000 دولار أميركي للسنة المالية 2019 (فترة 12 شهراً تنتهي في 30 حزيران/يونيو 2019) و59500 دولار في الربع التالي (الربع الأول من السنة المالية 2020، المنتهية في 30 أيلول/سبتمبر 2019) (Factcheck, 2019). ولم يتم الكشف عن مقدار المواد التي تم كشفها خلال تلك الفترات. كما وكشف موقع Snopes (الولايات المتحدة)، على الرغم من انسحابه من الشراكة مع فيس بوك في كانون الأول/ديسمبر 2018، عن حصوله على 406000 دولار أميركي في عام 2018 و100000 دولار في عام 2017 (من دون الإشارة إلى عدد الادعاءات المشبوهة التي تم كشفها). ووفقاً لعقود جهات التحقُّق من الوقائع الموصوفة لهيئة الإذاعة البريطانية BBC، يدفع فيس بوك رسوماً ثابتة لكل مقالة توضيحية، وهي في الولايات المتحدة حوالي 800 دولار (600 جنيه إسترليني) (Lee, 2019b).

تقييم مبادرة فيس بوك

لقد أجرى موقع Poynter.org دراسةً استقصائيةً شملت 19 منظمة تعمل بالشراكة مع فيس بوك (Funke & Mantzarlis, 2018a) ذكرت مجموعةً من الأسباب دفعتها إلى الانضمام إلى الشبكة. وقد انضمت مبادرة Full Fact البريطانية المشار إليها أعلاه، إلى برنامج فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة في أواخر العام 2018 ونشرت تقريراً تقييمياً (Full Fact, 2019) بعد ستة أشهر من توقيعها العقد. ولا يزال هذا التقييم يعدّ الأكثر تفصيلاً حول أداء البرنامج. وفيما تعتبر Full Fact أنّ برنامج فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة هو «جدير بالاهتمام»، وقد تكون شركات التواصل عبر الإنترنت الأخرى «بحاجة» إليه أيضاً، فقد أثارت المبادرة بعض المسائل والتوصيات المهمة، مثل: حاجة فيس بوك إلى التالي:

1. «تضمين المحتويات الموجودة على انستغرام بالكامل» في واجهة الويب لتوفير سلسلة من المحتويات المشبوهة لجهات التحقُّق من الوقائع[50].
2. «تطوير المزيد من الأدوات لتمكين جهات التحقُّق من الوقائع من البحث عن محتويات مشابهة وإبرازها» لمعالجة «تكرار نمط المعلومات المُضلّلة عبر الإنترنت» وتجنب معالجة جزء صغير من المشكلة فقط.
3. توفير المزيد من البيانات (لجهات التحقُّق من الوقائع) حول «مشاركة المحتويات التي يُحَذّر من زيفها عبر الزمن».
4. «إعطاء المزيد من البيانات إلى جهات التحقُّق من الوقائع حول مدى انتشار عمليات التحقُّق من الوقائع» لتقييم قيمة العمل الذي يتم في إطار البرنامج.

وتعتبر Full Fact أنّ نظام التصنيف الداخلي لفيس بوك غير مناسب» لغرض التحقُّق من الوقائع، حيث يصنّف المعلومات على أنها إما معلومات خاطئة أو مزيج من المعلومات الخاطئة والصحيحة أو عناوين خاطئة أو معلومات صحيحة أو غير خاضعة للتقييم، أو تهكُّم أو رأي أو مزاح أو معلومات غير مصنفة.

[50] منذ نشر تقرير Full Fact، أصبحت المحتويات الموجودة على انستغرام خاضعةً الآن للتحقق من الوقائع أيضاً، كما هو مذكور أعلاه.

ويستخدم فيس بوك شارات التصنيف الثلاث الأولى للحد من نشر المُحتوى وإخطار المستخدمين أن هذا المُحتوى قد خضع لعملية تحقق من الوقائع. واشتكت Full Fact من أن شارة «المزيج» غير كافية وتُفرِط في العقاب، حيث يتم وضعها على المُحتوى الذي يعتبر مزيجاً من المعلومات الصحيحة والخاطئة وتستخدم للادعاءات غير المثبتة؛ وبالتالي يتم الحد من نشر هذه المحتويات.

ورداً على بيان «مارك زوكربيرغ» أمام الكونغرس الأمريكي والذي توقّع فيه مزيداً من التحول نحو مقاربة تسمح بالإبلاغ عن مزيد من المحتويات بشكل مسبق من خلال أدوات تعتمد على الذكاء الاصطناعي (Zuckerberg, 2018)، صرّحت Full Fact إنها ترحّب ببيان أوضح من قبل الشركة حول «السبل المحتملة لتطوير أدوات التعلم الآلي» استناداً إلى بيانات شراكة التحقُّق من الوقائع من قبل أطراف ثالثة.

وعموماً، وفقاً لـ«مسح بوينتر» المذكور أعلاه، واستناداً إلى أهدافها الخاصة، يبدو أن جهات التحقُّق من الوقائع راضية بشكل مقبول عن الشراكة مع فيس بوك والأجر التي تتقاضاه مقابل عملها. ويبدو أنّ المسألة الأكثر إثارةً لمخاوف شركاء فيس بوك، وفقاً لاستطلاع «بوينتر»، هي أنّ الشركة لا تخبر الجمهور بما فيه الكفاية عن طريقة عمل الشراكة. كذلك، أظهر المسح أنّه لم يتم إثبات فعالية المبادرة لفائدة قدرتها على الحد من وجود المعلومات المُضلّلة على فيس بوك.

وقد خلُصت التحقيقات التي أجرتها Buzzfeed إلى أنّ «فيس بوك لا يزال موطناً للأخبار الزائفة واسعة الانتشار» (Silverman et al., 2017; Silverman & Pham, 2018). ولكن من جهةٍ أخرى، ظهرت بعض الأدلة التي تشير إلى انخفاض التفاعل مع المعلومات المُضلّلة على فيس بوك بعد الانتخابات الرئاسية الأميركية في العام 2016، حيث خلُصت ثلاث دراسات إلى أن هذا قد يعزى جزئياً إلى التدخلات الهادفة إلى التحقُّق من الوقائع. وخلصت دراسة أجراها باحثون في جامعة نيويورك وجامعة ستانفورد إلى أن التفاعل (من مشاركات وإعجابات وتعليقات) مع «الأخبار الزائفة» على فيس بوك انخفض من ذروة بلغت حوالي 200 مليون شهرياً في نهاية العام 2016 إلى حوالي 70 مليون شهرياً في تموز/يوليو 2018 (Allcott et al., 2018). وأشار الباحثون إلى أنه «... من الممكن أن تكون الجهود التي بذلها فيس بوك بعد انتخابات 2016 للحد من نشر المعلومات المُضلّلة قد أدت إلى تأثير فعلي». كما وقد انخفضت نسبة المعلومات المغلوطة إلى المعلومات المُضلّلة الموجودة على كل من فيس بوك وتويتر «بشكل حاد» وفقاً للدراسة، «... من حوالي 45:1 خلال انتخابات الولايات المتحدة في عام 2016 إلى حوالي 15:1 بعد عامين». ومع ذلك، وفقاً لهذا التقرير البحثي، لا يزال فيس بوك جهةً ناقلةً للتضليل أكثر بكثير من تويتر.

وقد أدخلت دراسة أكاديمية أخرى من جامعة ميشيغن ما يُسمَّى بـ«حاصل إيفي» "Iffy Quotient" لوصف المواقع التي تقوم في كثير من الأحيان بنشر معلومات مضللة (Resnick et al., 2019). وخلصت الدراسة إلى أن أداء فيس بوك وتويتر كان سيئاً خلال موسم انتخابات 2016، حيث تضاعف نشر المعلومات من مواقع مشكوك فيها مقارنةً بالمعدل الذي سجله في وقت سابق من ذلك العام. «ومع ذلك، سجّل فيس بوك انخفاضاً طويل الأجل في «حاصل إيفي» منذ آذار/مارس 2017»، كما أشار المؤلفون.

في مزيد من البحوث، قامت مبادرة Décodeurs، وهي عملية للتحقق من الوقائع تابعة للصحيفة الفرنسية لوموند Le Monde، بتحليل 630 موقعاً فرنسياً في عام 2018 بمساعدة ملحق المتصفح المسمى Decodex الخاص بها والذي يحذّر متصفحي الويب عند وصولهم إلى موقع إخباري مشكوك فيه أو مصدر معلومات

مضللة معروف من نوع آخر. وخلُصت المبادرة إلى أنّ التفاعل مع المواقع المشكوك فيها؛ المتدنية الصحة، والانتشار الواسع للأخبار الكاذبة انخفضا بشكل ملحوظ على فيس بوك (Sénécat, 2018).

وقد ألقى فيس بوك الضوء على هذه الدراسات في منشور مُدوَّن في عام 2018 يفيد فيه بأن هذه الدراسات تمثل دليلاً على أن «... الحجم الإجمالي للأخبار الخاطئة على فيس بوك يتجه نحو الانخفاض» (Lyons, 2018b). وفي الآونة الأخيرة، أثار إعلان من فيس بوك جدلاً حول سياسة الشركة في ما يتعلق بالتحقُّق من وقائع الإعلانات السياسية. فقد قررت الشركة أنها لن ترسل المحتويات العضوية أو الإعلانات من السياسيين أو من الجهات التابعة لهم إلى أطراف ثالثة شريكة للتحقق منها (Clegg, 2019).

وفي أوائل عام 2019، بعد بضعة أشهر من انسحاب ABC News (الولايات المتحدة) من برنامج فيس بوك للتحقق من الوقائع، قرر الموقع الأميركي Snopes العامل على مُكافحة الخدع؛ الانسحاب من البرنامج على الرغم من أنّ 33% من دخله في عام 2018 كان مصدره الشراكة مع فيس بوك (Green & Mikkelson, 2019). وفي نهاية تشرين الثاني/نوفمبر 2019، أعلنت جهة التحقُّق من الوقائع الهولندية Nu.nl انسحابها من البرنامج وسط جدل حول إعفاء فئات معينة من الإعلانات السياسية من التحقُّق من الوقائع من قبل الشركاء (Hern, 2019a).[51]

وتقضي سياسة فيس بوك عموماً بإعفاء الخطاب السياسي من التحقُّق من الوقائع، أي المنشورات والإعلانات من قبل السياسيين والأحزاب السياسية والجهات التابعة لها. ومع ذلك، تنص السياسة على أن التحقُّق من الوقائع يمكن أن يشمل «منظمات مثل Super PACs أو منظمات المناصرة غير التابعة للمرشحين». كما وتنص على ما يلي:

> ”عندما يشارك سياسي مُحتوى معيناً – أي رابطاً لمقال ما أو مقطع فيديو أو صورةً قام بتصميمها شخص آخر كان قد تم كشف زيفه على فيس بوك – سنقوم بالتقليل من أهمية هذا المُحتوى وعرض تحذير ورفض إدراج هذا المُحتوى في الإعلانات. وهذا يختلف عن الادعاءات أو البيانات التي تصدر عن هذا السياسي. فإذا تقدم السياسي مباشرةً بادعاء ما على صفحته، في إعلان ما أو على موقعه على الإنترنت، يُعتبر ذلك خطاباً مباشراً لا يخضع لبرنامج التحقُّق من الوقائع من قبل أطراف ثالثة حتى لو تم كشف زيف مضمون هذا الادعاء في مكان آخر.[52]“

ومع ذلك، فيما كانت توضع اللمسات الأخيرة على هذه الدراسة في تموز/يوليو 2020، أزال فيس بوك منشوراً للرئيس الأميركي السّابق «دونالد ترامب» لأول مرة، وذلك لانتهاكه سياسة فيس بوك المتعلقة بجائحة فيروس كورونا. وتضمنت المقالة مقطعاً يدعي فيه ترامب أنّ الأطفال هم «تقريباً محصنون كلياً» ضد الفيروس. ووفقاً لمتحدث باسم الشركة: «يتضمن هذا الفيديو ادعاءات كاذبة بأنّ مجموعة من الأشخاص محصنة ضد فيروس كورونا، وهو انتهاك لسياستنا حول المعلومات المُضلّلة والضارة المتعلقة بالفيروس». (BBC, 2020d; Carrie Wong 2020)

[51] يُرجَى مراجعة القسم 7.1 من الفصل السابع.

[52] https://www.facebook.com/business/help/182222309230722

وفـي مـا يتعلـق باسـتقلالية عمليـة التحريـر، أصـدرت Politifact العامـة فـي أواخـر عـام 2018 بيانـاً علـى تويتـر أيدتـه FactCheck.org (الولايـات المتحـدة) وAgência Lupa (البرازيـل) وTeyit.org (تركيـا) لنفـي تقريـر ورد فـي صحيفـة الغارديـان The Guardian يدعـي أنّ «فيـس بـوك يدفـع بالصحفييـن إلـى إعطـاء الأولويـة لكشـف المعلومـات المُضلّلـة التـي تؤثـر علـى الجهـات التـي تضـع إعلانـات علـى فيـس بـوك» (PolitiFact, 2018).

ووفقـاً لبعـض الـوكالات الإخباريـة التـي تقـوم بعمليـات كشـف الزيـف كأعضـاء فـي البرنامـج، لا يمنعهـا فيـس بـوك مـن التحقُّـق مـن وقائـع منشـورات السياسـيين والأحـزاب السياسـية (بمـا فـي ذلـك الإعلانـات السياسـية) ولكنـه لا يدفـع لهـا للقيـام بهـذا العمـل، كمـا ولا يتـم وضـع شـارة تصنيـف علـى هـذه المنشـورات علـى المنصـة عندمـا يتضـح أنهـا خاطئـة أو مضللـة. فعلـى سـبيل المثـال، فـي عـام 2019، قامـت وكالـة الأنبـاء الفرنسـية بالتحقُّـق مـن تصريحـات زعيمـة اليميـن المتطـرف الفرنسـي ماريـان لوبـان خمـس مـرات علـى مدونتهـا Factuel (Daudin, 2019) وعلـى حسـابها علـى فيـس بـوك، ولكـن هـذا لـم ينعكـس فـي الجـو العـام لفيـس بـوك بسـبب سياسـته التـي تحـد مـن التحقُّـق مـن وقائـع المحتويـات السياسـية.

أمـا المنشـورات التـي تُعتبـر آراءً، فهـي مجـال آخـر غيـر متفـق عليـه. فتنـص سياسـة فيـس بـوك علـى أن الـرأي «عمومـاً غيـر مؤهـل ليخضـع للتقييـم مـن قبـل جهـات التحقُّـق مـن الوقائـع. ويشـمل ذلـك المحتويـات التـي تناصـر أفـكاراً معيّنـةً وتسـتخلص الاسـتنتاجات اسـتناداً إلـى تفسـير معيّـن للوقائـع والبيانـات، وتطلـع الجمهـور علـى رأي الكاتـب أو المسـاهم فـي حـدث أو قضيـة مـا»[53]. وتتضمـن السياسـة تحذيـراً بـأنّ « ... المُحتـوى الـذي يُقـدم كـرأي ولكنـه مبنـي علـى معلومـات كاذبـة أساسـية قـد يبقـى خاضعـاً للتقييـم». ومـع ذلـك، فـإنّ هـذه السياسـة تعانـي مـن ثغـرات أدت إلـى إثـارة انتقـادات وجـدل واسـع حولهـا. ومـن الأمثلـة علـى ذلـك حالـة تتعلّـق بمقـال رأي مـن مجموعـة تنكـر وجـود تغيُّـر المنـاخ اسـتناداً إلـى تأكيـدات كاذبـة ومضللـة بشـأن علـوم المنـاخ. فـي هـذه الحالـة، قـام شـريك فيـس بـوك للتحقـق مـن الوقائـع المناخيـة Climate Feedback بتصنيـف المقالـة علـى أنهـا «خاطئـة»[54]، ولكـن بعـد التمـاس مُقـدَّم مـن المجموعـة، أزال فيـس بـوك التصنيـف علـى أسـاس أن المقـال كان «رأيـاً» ولا يخضـع للتحقـق مـن الوقائـع (Penney, 2020; Pastermak 2020). وفـي مثـال آخـر، أزال فيـس بـوك تصنيـف «غيـر صحيـح» كانـت قـد وضعتـه جهـة تحقـق مـن الوقائـع علـى شـريط فيديـو نُشـر علـى صفحـة أحـد الناشـطين المناهضيـن للإجهـاض علـى فيـس بـوك يدّعـي أنّ الإجهـاض لـم يكـن ضـرورةً طبيـةً قـط وذلـك بعـد شـكاوى متعـددة مـن مُشـرّعين مُحافظيـن (Grossman & Schickler, 2019). علـى الرغـم مـن أنّ الشـبكة الدوليـة للتحقـق مـن الوقائـع قـد حققـت فـي تقريـر جهـة التحقُّـق ووجـدت فـي أيلول/سـبتمبر 2019 أن الادعـاء الـوارد فـي شـريط الفيديـو هـو فعـلاً غيـر صحيـح[55]، إلا أنّ الفيديـو كان لا يـزال ينتشـر علـى فيـس بـوك بعد مرور عام من دون أي شارة تصنيف لوقائعه[56].

وقـد دفـع هـذا المثـال الأخيـر بشـكل خـاص المشـرعين فـي «مجلـس اللُّـوردات» فـي المملكـة المتحـدة إلى كتابـة مـا يلـي فـي تقريرهـم حـول «التحقيـق فـي التكنولوجيـا الرقميـة وإعـادة بنـاء الثقـة» أو Digital Technology

53 https://www.facebook.com/business/help/315131736305613?recommended_by=297022994952764

54 https://climatefeedback.org/evaluation/article-by-michael-shellenberger-mixes-accurate-andinaccurate-claims-in-support-of-a-misleading-and-overly-simplistic-argumentation-about-climate-change/

55 https://www.poynter.org/fact-checking/2019/the-ifcn-concludes-investigation-about-sciencefeedback/

56 https://www.facebook.com/youngamericasfoundation/videos/2113086642330235; https://www.facebook.com/youngamericasfoundation/videos/2113086642330235

and the Resurrection of Trust: «لم يتضـح وجـود أي شـكوك فعلية حـول دقـة عملية التحقُّق التي اسـتندت أيضـاً إلـى مراجعـة ثانوية مسـتقلة... وهـذا يشـير إلـى أنّ موقـف فيـس بـوك هدفـه تجنـب الضغـط السياسـي أكثـر منـه الحفـاظ علـى الحـوار الديمقراطـي». (House of Lords, 2020).

وسـيجري تنـاول الجـدل المتعلّـق بالإعلانـات السياسـية علـى فيـس بـوك (أي بشـأن سياسـة فيـس بـوك المتعلّقـة بالتحقُّق مـن الوقائـع المشـار إليها أعـلاه) فـي القسـم 5.3 المتعلق بالاسـتجابات الانتخابيـة، وفـي القسـم 7.1 الـذي يركّـز علـى الاسـتجابات الأخلاقيـة والمعياريـة.

ب. الاستجابات الإقليمية

منظمة أفريكا-تشيك AfricaCheck [57]

لقـد كانـت منظمـة AfricaCheck المحـرك الرئيـس لعمليـة التحقُّـق مـن الوقائـع فـي أفريقيـا. وهـي منظمـة غيـر ربحيـة أُنشـئت فـي عـام 2012 وهدفهـا تعزيـز صحـة المعلومـات المتداولـة فـي النقـاش العـام وفـي وسـائل الإعـلام الإخباريـة فـي أفريقيـا. وكان الهـدف هـو رفـع نوعيـة المعلومـات المتاحـة للمجتمـع فـي جميـع أنحـاء القـارة. ومنظمـة AfricaCheck مـن ابتـكار مؤسسـة وكالـة الأنبـاء الفرنسـية التـي تُعـد الـذراع غيـر الربحـي لوكالـة الأنبـاء الفرنسـية والتـي تسـعى إلـى تطويـر وسـائل الإعـلام. واليـوم AfricaCheck هـي منظمـة مسـتقلة لهـا مكاتـب فـي جوهانسـبرغ (جنـوب أفريقيـا) ونيروبـي (كينيـا) ولاغـوس (نيجيريـا) وداكار (السـنغال). وهـي تصـدر تقاريـر باللُّغتيـن الإنكليزيـة والفرنسـية وتقـوم بالتحقُّـق مـن الادعـاءات الصـادرة عـن شـخصيات ومؤسسـات عامـة ووسـائل إعـلام عبـر أفضـل الأدلـة المتاحـة. وقـد قامـت منـذ العـام 2012، بالتحقُّـق مـن أكثـر مـن 1500 ادعـاء حـول موضوعـات عديـدة تمتـد مـن الجريمـة والصراعـات العرقيـة فـي جنـوب أفريقيـا، إلـى أعـداد السـكان فـي نيجيريـا، و«علاجـات» صحيـة زائفـة فـي بلـدان أفريقيـة مختلفـة.

ويتـم نشـر عمـل Africa Check ومناقشـته فـي وسـائل الإعـلام الإخباريـة فـي جميـع أنحـاء القـارة. ومقرهـا الرئيسـي فـي قسـم الصحافـة فـي جامعـة ويتواترسـراند Witwatersrand فـي جوهانسـبرغ، جنـوب أفريقيـا، فـي حيـن أنّ موقـع اللغـة الفرنسـية يديـره فريـق مقـره فـي كليـة EJICOM للصحافـة فـي داكار، السـنغال، منـذ عـام 2015. وتعتمـد Africa Check علـى قُرائهـا لتحديـد الادعـاءات التـي يريـدون التحقُّـق منهـا، كمـا أنهـا تمكّـن وتشـجّع الصحفييـن الآخريـن علـى التحقُّـق مـن الادعـاءات بأنفسـهم بمسـاعدة قسـم التحقُّـق مـن الوقائـع والتقاريـر البحثيـة وخدمـات التعليـم.

منـذ إنشـائها، تلقـت Africa Check الدعـم التمويلـي مـن «مؤسسـة وكالـة الأنبـاء الفرنسـية» و«جامعـة آغـا-خـان» وقسـم الصحافـة فـي «جامعـة ويتواترسـراند» ومدرسـة EJICOM للصحافـة، بالإضافـة إلـى منـح مـن قائمـة طويلـة مـن المؤسسـات الخيريـة بمـا فـي ذلـك مبـادرة الإعـلام الأفريقـي والتحـدي الأفريقـي للابتـكار فـي مجـال الأخبـار، والمعهـد الدولـي للصحافـة، وغوغـل، ومؤسسـة كونـراد أدينـاور، وصنـدوق الألفيـة الائتمانـي، وLuminate، ومؤسسـات المجتمـع المفتـوح، ومؤسسـة Shuttleworth، ومؤسسـة بيـل وميلينـدا غايتـس، ومؤسسـة Raith، وبنـك سـتاندرد، وأبسـا وCode for Africa. وقـد سـجلت قاعـدة بيانـات جامعـة ديـوك للتحقـق مـن الوقائـع 17 منظمـةً فعّالـةً للتحقـق مـن الوقائـع فـي أفريقيـا فـي أوائـل عـام 2020.

[57] https://africacheck.org/

أمريكا اللاتينية: تأثير منظمة تشيكويدو Chequeado

في أمريكا اللاتينية، لعبت منظمة تشيكويدو Chequeado الأرجنتينية[58] دوراً بارزاً في مجتمع التحقُّق من الوقائع منذ إنشائها في عام 2010. وقد برزت العديد من المبادرات الجديدة في المنطقة منذ عام 2014، معظمها في مجال الصحافة، وذلك بفضل مساعدتها وتأثيرها، وخاصةً في مجال منهجيات التحقُّق من الوقائع. وفي عام 2019، قامت تشيكويدو بالتنسيق مع وكالة الأنباء الفرنسية في مشروع اسمه ريفيرسو Reverso project يُعنى بالتحقُّق من الوقائع خلال الحملة الانتخابية الرئاسية في الأرجنتين.

وتُسجل قاعدة بيانات التحقُّق من الوقائع في جامعة ديوك 16 منظمةً في أمريكا الجنوبية، وثمان في أمريكا الوسطى باللُّغة الإسبانية، و10 بالبرتغالية (في أوروبا، هناك 6 منظمات في إسبانيا ومنظمتان في البرتغال). وفي عام 2014، دعت تشيكويدو المنظمات الإقليمية الأخرى العاملة في التحقُّق من الوقائع إلى اجتماع في «بوينس آيرس» لإطلاق شبكة جديدة تُدعَى «LatamChequea» مصممة لتبادل أفضل الممارسات. ومنذ ذلك الحين، تعقد الشبكة الإقليمية مؤتمراً نصف سنوي في «بوينس آيرس». وتعقد أيضاً اجتماعات افتراضية شهرية بين منظمات التحقُّق من الوقائع تشمل أيضاً عدداً من علماء الاجتماع.

وتقوم بدعم Chequeado مالياً مؤسسة، La Voz Pública، كما وتنشط Chequeado في التعاون البحثي مع الأكاديميين حيث يلتحق المحققون في الوقائع لمدة أسبوع أو أسبوعين بغرفة أخبار Chequado عبر منح دراسية تُقدّمها الشبكة الدولية للتحقق من الوقائع.

أوروبا: سوما

إنّ المرصد الاجتماعي لتحليل المعلومات المُضلّلة ووسائل التواصل الاجتماعي (SOMA) مموّل من قبل المفوضية الأوروبية، وهدفه تنظيم جهات التحقُّق من الوقائع في أوروبا كجزء من جهد عابر للدول في أوروبا لإعادة بناء الثقة في الصحافة، وتقديم الدعم للمجتمع المُتنامي من المنظمات الإعلامية، والجهات المحققة في الوقائع والأكاديميين والمنظمات غير الحكومية وصانعي السياسات الذين يتصدون للمعلومات المُضلّلة. ووفقاً لمنصة Truly Media، انضمت نحو 40 منظمةً أوروبيةً رسمياً إلى هذا المرصد في السنة الأولى من عمله[59]. وقد نشر هذا المرصد الأوروبي العديد من التحقيقات والتوصيات المتعلقة بالمعلومات المُضلّلة حول جائحة فيروس كورونا. ومن المقرر أن يُتبع في السنوات القادمة بمرصد جديد مُسمى EDMO (المرصد الأوروبي للإعلام الرقمي)[60].

الدول العربية

في الدول العربية، لم يأخذ بعد التعاون بين مبادرات التحقُّق من الوقائع طابعاً رسمياً، ولكن جهات التحقُّق من الوقائع في المنطقة مترابطة بين بعضها وتتعاون في مجال التدريب وتجتمع في مؤتمرات على غِرار منتدى الإسكندرية للإعلام في مصر؛ الذي ركّز على التحقُّق من الوقائع، والتّضليل، والدراية والتدريب الإعلاميين في نُسخِه الثلاث الأخيرة التي عُقدت في الإسكندرية (2017–2018) والقاهرة (2019).

---

58 https://chequeado.com/

59 https://www.truly.media/

60 https://edmo.eu/

وعلى الصعيد الإقليمي، من المبادرات البارزة التي أُطلقت في تلك الآونة، مبادرة «فتبيّنوا» Fatabyyno ومقرها الأردن[61]. وقد تم إطلاقها في عام 2014، وهي تسعى إلى رصد وكشف المعلومات المُضلّلة في ثماني عشرة دولة في الشرق الأوسط وشمال أفريقيا. ومن المبادرات الأخرى «دا بجد» التي أُطلقت في مصر في عام 2013، وكذلك مُبادرة «ما تصدقش»[62] و«فالصو»[63]. [64] وهناك أيضاً مبادرتان اسمهما «فالصو» تعملان في مجال التحقُّق من الوقائع في ليبيا[65] وتونس[66].

وفي الجمهورية العربية السورية، منصة «تأكّد» المعروفة أيضاً باسم Verify Syria [67] هي موقع على شبكة الإنترنت لرصد وكشف المعلومات المُضلّلة، ينشر بثلاث لغات هي العربية والإنكليزية والتركية. أما عملية التحقُّق من الوقائع لوكالة الأنباء الفرنسية[68] فهي تغطي مختلف دول الشرق الأوسط وشمال أفريقيا ومقرها لبنان؛ وهي نتيجة تعاون مع جهة التحقُّق من الوقائع المعروفة «بالعسَّاس»[69].

ج. بعض الاستجابات الوطنية الأخرى

يتناول هذا القسم الفرعي مبادرات وطنية محددة وجديرة بالذكر في مجال الرصد والتحقُّق من الوقائع. في الولايات المتحدة وأوروبا، يرتبط تاريخ وتطور التحقُّق من الوقائع بالحملات الانتخابية والتحقُّق من الادعاءات السياسية. ولذلك، فإنّ تفصيل معظم هذه الاستجابات يرد في القسم 5.3.

الهند

في الهند، تطوّر واتساب المملوك لشركة فيس بوك ليصبح من قنوات التّضليل الرئيسية (Kajimoto & Stanley, 2019). وقد لوحظ أن ظاهرة استخدام السياسيين لوسائل التواصل الاجتماعي للنفاذ المباشر إلى الجماهير، متجاوزين وسائل الإعلام التقليدية التي تلعب دور الحارس، قد بدأت بالبروز منذ عام 2014، وقد ساعدت على انتشار المعلومات المُضلّلة داخل الشبكات الاجتماعية عبر الإنترنت (Kajimoto & Stanley, 2019).

وتعمل خمس عشرة منظمةً للتحقق من الوقائع في الهند وفقاً لقاعدة بيانات جامعة ديوك؛ ثمانية منها أعضاء في شبكة فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة. ومع ذلك، فإن هذه المبادرات هي في الغالب من الأفراد والمنظمات أو الفرق الصغيرة (مثل: المحققين في الوقائع لصحيفة Times of India). وجميعها تم إنشاؤه منذ عام 2013. وهي تشمل Factcheck وSM Hoax Slaye وBoomlive للصحافة

61 https://fatabyyano.net/and https://www.facebook.com/Fatabyyano/
62 https://www.facebook.com/matsda2sh/
63 https://www.falsoo.com/
64 https://dabegad.com/
65 https://falso.ly/
66 https://www.facebook.com/falso.tn/
67 https://www.verify-sy.com/
68 https://factual.afp.com/ar
69 https://twitter.com/el3asas

الاستقصائية التي تحوّلت إلى التحقُّق من الوقائع في عام 2016.[70][71] وجزء كبير من المعلومات المُضلّلة التي كشفت هذه الجهات عنها هو سياسي، أو محلي أو مصدره توترات جيوسياسية مرتبطة بالحدود والجغرافيا بصفةٍ خاصّة.

إندونيسيا

في إندونيسيا، تكافح منظمة «مافيندو» Mafindo غير الحكومية للتصدي المعلومات المُضلّلة منذ العام 2015 من خلال مجموعة مناهضة للتشهير والخداع[72] على فيس بوك، ومن خلال مجموعة كشف الخدع على واتساب[73]، ومن خلال وصلة على متصفح غوغل كروم، وموقع على شبكة الإنترنت[74] يستخدم شعارها «عكس الخدع». وبعد مافيندو، تم إطلاق خمس مبادرات أخرى لكشف المعلومات المُضلّلة في إندونيسيا، معظمها من قبل منظمات إخبارية، ست منهما جزء من شبكة فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة. ومبادرة أخرى ذكرها الباحثون (Kajimoto & Stanley, 2019)، وهي «كلاريفيكازيهوكس» Klarifikasihoax توقفت عن العمل منذ عام 2017.

الفلبين

في الفلبين، تؤدي حملات التّضليل إلى خلق مجال عام مليء بتلوّث المعلومات، وبالتالي، مليء بالتحريض السام والاستقطاب، وذلك منذ الانتخابات الرئاسية لعام 2016. وكما ورد في إحدى مطبوعات اليونسكو (Posetti 2017)، ازدادت قوة اللّجان الإلكترونية على الإنترنت. وهي تستخدم الشبكات «الدُّمَى» مع ما يتأتى عن ذلك من عواقب محتملة على المدى الطويل على الديمقراطية والانتخابات (مراجعة Ong & Cabañes, 2018; Ressa, 2016). ومع ذلك، تقوم أربع منظمات للتحقق من الوقائع برصد المعلومات المُضلّلة والادعاءات السياسية، بما في ذلك Vera files[75] ورابلر Rappler[76]، وثلاثة منها أعضاء في شبكة فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة.

جمهورية كوريا

وفي جمهورية كوريا، انتشرت الشائعات والدعاية الحزبية والمعلومات المُضلّلة على تطبيقات المراسلة مثل KakaoTalk أو Naver Band، وكذلك على مواقع التواصل الاجتماعي، لا سيما أثناء الانتخابات. ومن إحدى المبادرات الرئيسية التي تم إعدادها للانتخابات الرئاسية لعام 2017 مبادرة SNU Factcheck[77]، التي أطلقتها «جامعة سيول الوطنية» لجمع 26 وسيلةً إخباريةً للتحقق من المعلومات غير المؤكدة. ولا زالت واحدة من المنظمات الخمس العاملة في مجال التحقُّق من الوقائع في البلاد. وليس في شبكة

....

70 https://smhoaxslayer.com/about/

71 https://www.factchecker.in/about-us/

72 https://www.facebook.com/groups/fafhh

73 https://mafindo.gitbook.io/whatsapp-hoax-buster/

74 https://turnbackhoax.id/

75 https://verafiles.org/

76 https://www.rappler.com/newsbreak/fact-check

77 http://factcheck.snu.ac.kr/

فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة أعضاء من كوريا، وذلك مردّه أساساً إلى أنّ Naver و Daum المحليتين هما البوابتان الأكثر شعبيةً على الإنترنت، وهما تتبعان سياسةً تطلبان فيها من مزودي المُحتوى الإخباري الخضوع لعملية تقييم؛ مما يُصعّب على مروجي المعلومات المُضلّلة وضع المُحتوى على هاتين البوابتين.

الولايات المتحدة – Snopes

Snopes هي من رُواد كشف المعلومات المُضلّلة والتحقُّق من الوقائع في الولايات المتحدة. ففي 1994، قام ديفيد ميكلسون بتأسيس موقع snopes.com[78] كهواية له في التحقيق في الأساطير الشعبية والخدع على المجموعة الإخبارية alt.folklore.urban على شبكة Usenet (وهي قناة حوار عالمية ظهرت في بدايات الإنترنت).

ومباشرة بعد الهجمات الإرهابية في 11 أيلول/سبتمبر 2001، بدأ مؤسسو snopes.com في كشف الشائعات والأكاذيب حول الهجمات[79] فقاموا بتقييم حوالي 176 أسطورةً وشائعةً بين 2001 و2011 (Aspray & Cortada, 2019). فكانت هذه نقطة تحول في حياة snopes.com التي تحوّلت من منصة لكشف الأساطير الشعبية كهواية، لتصبح تدريجياً من أهم المنظمات العاملة في مجال التحقُّق من الوقائع بصورة منهجيّة احترافيّة (Dean, 2017). وبين الأخبار العاجلة (مثل إعصار كاترينا في 2005) والانتخابات الرئاسية (في عام 2008 والشائعات المتداولة حول مكان ولادة الرئيس الأميركي الأسبق باراك أوباما حتى استطلاع الآراء في 2016)، ازداد جمهور snopes. com مما سمح للموقع بأن يصبح مشروعاً يحقق استدامته من خلال الإيرادات الإعلانية[80].

5.1.4 دراسة حالة بشأن الاستجابة: المعلومات المُضلّلة حول جائحة فيروس كورونا

صرّح المدير العام لمنظمة الصحة العالمية السيد «تيدروس أدهانوم غيبريسوس» بأن جائحة فيروس كورونا أرغمت البشرية على مُكافحة جائحة صحية و«أزمة معلومات مضللة» في الوقت نفسه (Zarocostas, 2020). وقد اتخذت كافة المنظمات التي تمت مراجعتها في هذا الفصل تدابير للاستجابة لجائحة فيروس كورونا من خلال محاور أو صفحات خاصة حول المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا.

على سبيل المثال، نشرت First Draft مجموعةً كاملةً من الموارد للصحافيين[81] على غرار أدوات وأدلة ومبادئ توجيهية حول الأخلاقيات ووحدة تعليمية عبر الإنترنت وقاعدة بيانات قابلة للبحث تتضمن كشفاً

78 https://www.snopes.com/

79 https://scholar.colorado.edu/concern/book_chapters/8049g572m

80 يُرجَى مراجعة ما ورد سابقاً في هذا الفصل حول الموضوع، والقسم 7.1 في الفصل السابع، الذي يفصّل دور Snopes كعضو في شبكة فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة.

81 https://firstdraftnews.org/long-form-article/coronavirus-resources-for-reporters/

للمعلومات المُضلّلة حول الفيروس، استناداً إلى أداتين للرصد: مُستكشف غوغل للتحقق من الوقائع[82]، وقاعدة بيانات تحالف الوقائع حول جائحة فيروس كورونا التابعة للشبكة الدولية للتحقق من الوقائع[83]. وقد تم إطلاق هذه المبادرة الأخيرة في كانون الثاني/يناير 2020 كحملة وسم «هاشتاغ» مزدوجة على تويتر #CoronaVirusFacts (باللغة الإنكليزية) و#DatosCoronaVirus (باللُّغة الإسبانية) لأعضاء الشبكة المشاركين، وذلك عندما كان الوباء لا يزال محصوراً في الصين، ولكن كان يجري استغلاله لأغراض التّضليل.

وأدت حملة الوسم «الهاشتاغ» إلى قاعدة بيانات تضم أكثر من 3000 عملية تحقق من الوقائع من 70 دولة و40 لغة (في نيسان/أبريل 2020). بعد ذلك[84]، سعى مشروع آخر بقيادة Science Feedback، برعاية مُبادرة أخبار Google، إلى توسيع قاعدة البيانات هذه وتضمينها جميع المواقع وعناوين URL التي تشارك معلومات مضللة حول جائحة فيروس كورونا.

ونشر مجتمع التحقُّق من الوقائع، من الشبكة الدولية للتحقق من الوقائع وبرنامج فيس بوك للتحقق من الوقائع من قبل أطراف ثالثة وغيرهما من الجهات، تقارير كثيرة لكشف المعلومات المُضلّلة حول تفشي فيروس كورونا، وقام بتسجيل حالات تضليل انتشرت في كُل ربوع العالم. ووفقاً لدراسة أجراها معهد رويترز لدراسة الصحافة (RISJ) استناداً إلى تحليل لعمليات التحقُّق من الوقائع باللُّغة الإنكليزية برعاية First Draft، تبيّن أنّ عدد عمليات التحقُّق من الوقائع ارتفع بأكثر من 900% بين كانون الثاني/يناير إلى آذار/مارس 2020. وبناءً على تحليل 225 عملية تحقق، وجد المعهد أنّ 59% من المعلومات المُضلّلة هي معلومات موضوعة بأشكال ثانية بينما 38% من هذه المعلومات هي معلومات مُلفّقة. (Brennan et al 2020)

وقد قامت بعض شركات التواصل عبر الإنترنت (مثل: فيس بوك[85]، ويوتيوب[86]، وانستغرام[87 88]، وواتساب[89]، وتويتر، ولينكدن[90]) باتخاذ إجراءات لربط مستخدميها بمعلومات موثوقة حول الجائحة من خلال إحالة أي استفسار عن الفيروس إلى المركز الرئيسي لمنظمة الصحة العالمية[91] وصفحة الكشف عن زيف المعلومات المُضلّلة التابعة لمنظمة الصحة العالمية[92]، أو لوزارات الصحة المحلية. وهي تقوم أيضاً بنقل التحذيرات الصادرة عن منظمة الصحة العالمية من خلال روبوتات الدردشة وتلك الصادرة عن السُّلطات المحلية عبر تطبيقات المراسلة[93]. كما وتنشر صفحات رسمية تتضمن معلومات مبنية

82 https://toolbox.google.com/factcheck/explorer

83 https://www.poynter.org/ifcn-covid-19-misinformation/

84 https://sciencefeedback.co/building-an-open-source-database-of-misinformation-sources-oncovid-19/

85 https://www.facebook.com/coronavirus_info/?page_source=coronavirus_hub_attachment&fref=mentions

86 https://www.youtube.com/watch?v=i352PxWf_3M

87 https://about.instagram.com/blog/announcements/coronavirus-keeping-people-safe-informedand-supported-on-instagram/

88 https://www.whatsapp.com/coronavirus

89 https://blog.twitter.com/en_us/topics/company/2020/covid-19.html

90 https://www.linkedin.com/feed/news/coronavirus-official-updates-4513283/

91 https://www.who.int/emergencies/diseases/novel-coronavirus-2019

92 https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters

93 https://www.messenger.com/coronavirus

على الوقائع. ويروّج البعض أيضاً لمنظمات التحقُّق من الوقائع التابعة للشبكة الدولية للتحقق من الوقائع[94] ويطلب من المستخدمين التحقُّق من الوقائع والامتناع عن مشاركة المعلومات في حال عدم التأكد من صحتها.

وقد أعلن كلّ من غوغل (Mantzarlis, 2020) وفيس بوك (Goldshlager & Watson, 2020) وواتساب (IFCN, 2020a) عن برامج صغيرة لتمويل جهات التحقُّق من الوقائع والمنظمات غير الربحية التي تحارب المعلومات المُضلّلة حول الجائحة في العديد من البلدان (IFCN, 2020b). وتم الإعلان عن ثلاثة عشر مشروعاً في العدد نفسه من البلدان في بداية نيسان/أبريل من خلال برنامج «منح للتحقق من الوقائع المتعلقة بجائحة فيروس كورونا» (IFCN, 2020c).

بالإضافة إلى ذلك، قررت شركات التواصل عبر الإنترنت «العمل معاً بشكل وثيق»[95] لمُكافحة الاحتيال والمعلومات المُضلّلة المرتبطة بالوباء. وقد بدأت العديد من الشركات بحجب الإعلانات التي تحاول الاستفادة من المعلومات المُضلّلة المتعلقة بالفيروس وإزالة المعلومات المُضلّلة التي قد تؤدي إلى ضرر جسدي. على سبيل المثال، في نيسان/أبريل، صرح فيس بوك أنه وضع 50 مليون شارة تحذيرية على منشورات وُضعت على المنصة استناداً إلى أكثر من 7500 مقالة من شركاء عاملين في التحقُّق من الوقائع[96]. كما وتقوم بعض الشركات بإزالة المنشورات من نوع المؤامرة، باستخدام أحكام سياسية تتعلق بتقييم المُحتوى بداعي إحداثه للضرر[97]. في حالة فيس بوك، فإن هذه المنشورات تزرع «التّضليل والشائعات التي لا يمكن التحقُّق منها، والتي تسهم في نشر خطر العنف الوشيك أو الأذى الجسدي»[98]. وقد صرح «زوكربيرغ»، الرئيس التنفيذي لشركة فيس بوك بأنّ التمييز بين المعلومات الصحيحة والخاطئة خلال جائحة ما أسهل من التمييز بين هذه المعلومات خلال الحملات السياسية (Smith, 2020a).

وليس من الممكن قياس مقدار المعلومات المُضلّلة حول جائحة فيروس كورونا التي قامت هذه الشركات بالتحقُّق منه لأنها عادةً لا توفر إحصاءات مفصلة حول مصادر وأرقام وأنواع الأمور التي تم فحصها، ولا بيانات عن تداول هذه المنشورات قبل اتخاذ أي إجراءات (UNESCO, 2020)

وعلى الرغم من جهود التحقُّق من الوقائع، حذّرت الوثائق الصادرة عن معهد الحوار الاستراتيجي (ISD, 2020a) في آذار/مارس ونيسان/أبريل 2020 من استغلال جائحة فيروس كورونا من قبل الشبكات المعادية للمهاجرين والكارهة للأجانب أو من قِبل اليمين المتطرف (ISD, 2020b)، خاصةً في المجموعات المغلقة على فيس بوك، وقنوات الدردشة على تطبيق واتساب، والشبكات الهامشية، مثل 4chan (Arthur, 2019)، وبلغات أخرى غير الإنكليزية.

94 https://faq.whatsapp.com/126787958113983

95 https://twitter.com/googlepubpolicy/status/1239706347769389056

96 https://about.fb.com/news/2020/04/covid-19-misinfo-update/

97 https://blog.twitter.com/en_us/topics/company/2020/covid-19.html

98 https://www.facebook.com/communitystandards/credible_violence/

6.1.4 كيف يتم تقييم استجابات الرصد والتحقُّق من الوقائع؟

يمكن تقييم التحقُّق من الوقائع لفائدة تحقيقه لأهدافه الفورية طويلة الأجل. ويعتمد ذلك على تقييم حجمه ومدى وصوله وحسن توقيته، من بين عوامل أخرى. ومع ذلك، لا يعد ذلك أمراً بسيطاً؛ حيث لا تزال البحوث المنشورة عن مدى وصول التحقُّق من الوقائع وتأثيره محدودة. كذلك، تحتفظ شركات الإنترنت بالكثير من البيانات ذات الصلة. وهذا يصعّب التقييم ويجعل الباحثين يقومون بالاستقراء من بيانات محدودة (مثل: مقاييس الوصول والتفاعل المرتبطة بتقارير كشف المعلومات المُضلّلة وفحص الوقائع التي تنشرها الوكالات الإخبارية)، ومن البحوث (مثل: الدراسات الإثنوغرافية والنفسية) حول استجابات المواطنين المتنوعة للمعلومات المُضلّلة والتدابير التصحيحية. بالإضافة إلى ذلك، من الصعب اختبار الافتراض الكامن القائل بأن الأدلة التي تم التحقُّق منها والفكر العقلاني يلعبان دوراً في مُكافحة المعلومات المُضلّلة بسبب الترابط المعقّد بين المعلومات المُضلّلة والعاطفة والهُويّة.

وخلال الفترة التي سبقت الانتخابات الرئاسية الأميركية لعام 2012، ظهرت مخاوف بشأن فعاليّة التحقُّق من الوقائع وقدرته على الحد من المعلومات المُضلّلة، لا سيما تلك المرتبطة بالخطاب السياسي. ولكن لا يزال هناك زخم كافٍ لمواصلة جهود التحقُّق من الوقائع، كما يتضح من تصريح أحد الباحثين: «إنّ تسمية وكشف السياسيين الذين يُضلّلون الجمهور مراراً وتكراراً قد يلحق أضراراً كبيرةً بسمعتهم مع مرور الوقت» (Nyhan, 2012).

وقد ورد في العديد من الدراسات أنّ تعزيز مشاركة المواطنين في الشؤون العامة، وزيادة التكلفة المتعلقة بالسمعة على السياسيين، وزيادة ثقة الجمهور بوسائط الإعلام، كلها أمور لديها آثار إيجابية على التحقُّق من الوقائع. ومع ذلك، فإن استعداد المواطنين لقبول التصحيحات التي تعزز وجهات نظرهم الخاصة أمر مهم. وتساعد النتائج العكسية أو «التأثير المُرتد» على نشر المعلومات المُضلّلة (أي إذا كانت عمليات التحقُّق من الوقائع تتعارض مع آراء المواطنين القائمة حول جهة أو قضية سياسية، فمن المرجح أن يتم رفضها على الرغم من صحتها)، خاصةً عندما يتم استعمال المعلومات المُضلّلة (والتحقُّق من الوقائع) كسلاح من قبل السياسيين أنفسهم لزيادة الاستقطاب.

ومن منظمات التحقُّق من الوقائع التي حاولت تقييم أثر عملها منظمة Chequeado الأرجنتينية. ففي مراجعة لـ6 دراسات أكاديمية لتقييم أثر التحقُّق من الوقائع في الولايات المتحدة، نظر الباحثون المكلفون من شبكة Chequeado في فعالية جهودهم وفي أثر التحقُّق من الوقائع على المواطنين والنخب السياسية والمنظمات الإعلامية (Pomares & Guzman, 2015) فوجدوا أن تعزيز مشاركة المواطنين في الشؤون العامة، وزيادة التكلفة المتعلقة بالسمعة لنشر السياسيين للأكاذيب، وتعزيز ثقة الجمهور في وسائل الإعلام هي آثار إيجابية للتحقق من الوقائع. ومن الاستجابات التقييمية التي اقترحها الباحثون بتكليف من Chequeado قياس قوة التحقُّق من الوقائع (على سبيل المثال، هل يصمد أمام الفحص الدقيق) ومدى وصوله. (Pomares & Guzman, 2015)

وأورد تقرير مشترك حديث من Africa Check وChequeado وFull Fact أنّ التصحيحات العامة للبيانات أو الإحصاءات المُضلّلة، ووقف الادعاءات الكاذبة من السياسيين، والإفصاح عن بيانات جديدة ذات مغزى، وتدريب الصحفيين في غرف التحرير القديمة على الحد من انتشار المعلومات المُضلّلة، وإشراك المسؤولين والجهود الرامية إلى تعزيز المُساءلة؛ كلها نتائج إيجابية للتحقق من الوقائع (Africa Check, Chequeado & Full Fact, 2020) ويمكن اعتبارها مؤشرات على الفعالية.

7.1.4 التحديات والفرص

إنّ حجم واختلاف أنواع المعلومات المُضلّلة يصعّبان عملية تحديد المشكلة ورصدها والإبلاغ عنها ولفت انتباه الجمهور إلى جميع مظاهرها وأبعادها. كما وتظهر فروقات دقيقة أساسية. فهناك مثلاً فرق بين إثبات زيف ادعاء ما وبين إثبات عدم وجود دليل عليه حالياً (ولكن قد يكون صحيحاً) وبين تبيان صحة ادعاء ما على أساس وقائع عندما يكون هناك سردية أو منظور أوسع يحشد ويجمع بين حقائق معينة ويقدمها مصحوبةً برأي وموقف وهوية معيّنة.

ومما يزيد من تعقيد الأمر هو تقييم الآثار المقصودة وغير المقصودة لتحديد المُحتوى الزائف ومنتجيه. غير أن إنتاج هذا النوع من التحاليل أمر حيوي لتطوير أو تعديل عمليات التحقُّق من الوقائع وغير ذلك من الاستجابات.

ويكمن التحدي الذي تواجهه جهات التحقُّق من الوقائع في تطبيق معايير موضوعية والعمل بشفافية في جميع البلدان واللُّغات وعلى نطاق واسع وبشكل يُحدث أثراً. وهذا ضروري لتزويد المجتمع بالمعلومات الضرورية التي تضمن فعالية مختلف الاستجابات على النحو الأمثل. ولكنّ تحقيق ذلك من الناحية العملية ليس سهلاً أبداً لا سيما في حالة اللغات غير المنتشرة عالمياً والبلدان الأصغر حجماً، التي غالباً ما لا يكون لديها منظمات محلية مستقلة معنية بالتحقُّق من الوقائع. لذلك، تعمل المنظمات الدولية للتحقق من الوقائع على سد هذه الفجوة، ولكنها تعتمد على ناطقين أصليين بلغة معيّنة، مما يحد من إمكانية التدقيق في عملهم وتحيزاتهم. وهذا يُسلّط الضوء على الحاجة إلى مقاربة مُحكَمة ومستقلة تسعى إلى «تقييم المُقيّمين» أو «التحقُّق من العاملين في التحقُّق من الوقائع».

وينبغي أيضاً أن يراعي التحقُّق من الوقائع المعايير الدولية لحُريّة التّعبير وغيرها من حقوق الإنسان، مثل احترام الخصوصية. كما ولابد من أن يعترف بأن مُحتوى معيّن (على غرار المجهول وبعض السرديات والآراء والفكاهة) غير قابل للتحقق. كذلك، يجب أن يرقى التحقُّق من الوقائع إلى مستوى قيم الشفافية وعدم التحزّب وتجنب الانتقائية.

ويتمثل التحدي الذي تواجهه منظمات التحقُّق من الوقائع في درء الهجمات القانونية عليها. وقد ظهرت مبادرة الدعم القانوني لجهات التحقُّق من الوقائع Factcheckers Legal Support Initiative لتقديم المشورة القانونية[99]. وهي تجمّع لمبادرة الدفاع القانوني عن وسائل الإعلام، والشبكة الدولية للتحقق من الوقائع، ولجنة المراسلين من أجل حُريّة الصحافة.

وتُتيح الأحداث الكبرى، مثل الانتخابات وحالات الطوارئ الصحية العامة، فرصةً تسمح للاستجابات المستقلة للرصد وتحديد المعلومات المُضلّلة إعادة التأكيد على قيمة الوقائع، وتشجيع البحث العام في ما يُعتبر مُحتوى موثوقاً يمكن مشاركته. فعلى سبيل المثال، يتيح تحديد المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا والتحقيق في الاستجابات على مر الزمن التقييم المستمر لفعالية شركات التواصل عبر الإنترنت في «تسطيح مُنحنى» «أزمة المعلومات المُضلّلة» (Proctor, 2020; Posetti & Bontcheva 2020a; Posetti & Bontcheva 2020b). كما أن استجابات تحديد المعلومات المُضلّلة أساسية لرصد تقاطع هذه المعلومات مع خطاب الكراهية

99 https://factcheckerlegalsupport.org/

المستخدم ضد النساء والأقليات والمهاجرين وغيرهم من المواطنين والمجتمعات الضعيفة. غير أنّ بضعاً من هذه الجماعات المستهدفة قد يلجأ أيضاً إلى أساليب التّضليل ممّا يجعل من الضروري فحص المُحتوى الذي ينتجه. ففي بعض الأحيان تقوم بعض المجموعات التي تقع ضحية المعلومات المُضلّلة باللُّجوء إلى الأساليب نفسها لدعم قضاياها. لذلك لابد من إخضاع المنشورات التي تنتجها إلى التحقُّق من الوقائع والتدقيق فيها.

ومن شأن ذلك أن يُعزّز أيضاً من استجابات تحديد المعلومات المُضلّلة. وقد تعهّد كل من واتساب (IFCN,2020a) وفيس بوك (Axelrod, 2020) وغوغل[100] وتويتر (Gadde, 2020) بتخصيص الأموال لمنظمات التحقُّق من الوقائع، ممّا يدلّ أيضاً على أنه يمكن القيام بالمزيد في هذا الخصوص. ولابد من دعم مستمر خلال الفترات المهمة كالانتخابات والأوبئة وما بعدها. ويتسم التحقُّق من الادعاءات المتعلقة باللقاحات وتغير المناخ بأهمية خاصة في المستقبل.

8.1.4 توصيات بشأن الاستجابات المتعلقة بالرصد والتحقُّق من الوقائع

إن التحديات والفرص المحددة أعلاه والحالة الراهنة للتحقق من الوقائع وكشف المعلومات المُضلّلة تؤدي إلى وضع التوصيات السياستية التالية للمؤسسات الدولية والإقليمية والحكومات وشركات التواصل عبر الإنترنت والمؤسسات والوكالات الإخبارية:

- إتاحة الموارد للتحقق المستقل من الوقائع، بما في ذلك تيسير التحقُّق من الوقائع في ما يتعلق بالمُحتوى السياسي والإعلانات السياسية.
- دعم مبدأ النفاذ إلى المعلومات ذات الصلة بالتحقُّق من الوقائع، لا سيما في ما يخص السُّلطات وشركات التواصل عبر الإنترنت، وذلك من أجل تعزيز الشفافية وتمكين منظمات التحقُّق من الوقائع من العمل بشكل أكثر دقةً وشفافية.
- التأكيد على أن نتائج التحقُّق من الوقائع هي مصادر معلومات جديرة بالثقة، ومفيدة للمواطنة، ووسائط الإعلام، وللتدخلات الهادفة إلى تعزيز الدراية الإعلامية والمعلوماتية.
- تعزيز البحوث متعددة التخصصات حول استجابات التحقُّق من الوقائع المُضلّلة.
- المساعدة في تطوير عمليات تحقق من الوقائع تتسم بالتعاون في جميع أنحاء العالم وذلك لتحقيق النفاذ إلى معلومات دقيقة وموثوق بها على الصعيد العالمي، وخاصةً بالشراكة مع المؤسسات الإخبارية.
- تعزيز القدرة على التحقُّق من الوقائع داخل المؤسسات الإخبارية من خلال تدريب متخصص ومشروعات تحريرية تسعى إلى دعم تقارير المُساءلة حول الجهات الفاعلة المختلفة من شركات وحكومات وجهات سياسية وحول إجراءات كل منها.
- تطبيق «التحقُّق من المُحقّقين» ووضع معايير دولية ومقاربة للمُساءلة تسمح باعتماد إجراءات تعيين وتقييم شفافة وموضوعية للأشخاص والمنظمات المُشاركة في التحقُّق من الوقائع، وبتقييم أدائها عبر الزمن (بما في ذلك شركات التواصل عبر الإنترنت التي تُسهّل وتُموّل فحص الوقائع على مواقعها).

[100] https://www.france24.com/en/20200402-google-boosts-support-for-checking-coronavirus-facts

2.4 الاستجابات الاستقصائية

المؤلفون: سام غريغوري، جولي بوزيتي، دينيس تيسو

يتناول هذا القسم مجموعة الكيانات التي تجري تحقيقات في المعلومات المُضلّلة (بدءً من تحقيقات الصحافة وصولاً إلى تحقيقات المجتمع المدني) ونواتج عملها.

وتركّز التقارير الاستقصائية عادةً على حملات وجهات فاعلة معينة، وتتعدى في تحقيقها مسألة تحديد صحة أو زيف المعلومات ومقدار هذا الزيف (التحقُّق من الوقائع). وقد تشمل وتتعدى مسألة تحديد ما إذا كان منشور معيّن قد خضع للتلاعب أو إلى سوء التصوير أو إلى تجريده من سياقه (كما في حالات وضع شارة تحديد المصدر)، أو ما إذا كان منشور معيّن أو وسيلة إعلامية معيّنة موثوق بها ومنصفة (على سبيل المثال، وضع شارة تقييم للمصداقية). ومن المرجح أن ترصد (قدر الإمكان) حالات وموضوعات وانتشار المعلومات المُضلّلة المعينة. وعندما يتم نشرها بالكامل، توفر هذه التقارير رؤىً ثاقبة عن حركيّة حملات التّضليل، بما في ذلك عن عناصر مهمة على غِرار الشبكات التي تُجريها وأهدافها والوسائط والأساليب المستخدمة والميزانيات المتاحة. كما وتمسح بتحديد الإسناد والنية. للحصول على أمثلة على تصنيفات وفهرسة الحملات، الرجاء مراجعة Bradshaw & Howard (2018) وBrooking et al. (2020)
وتهدف هذه التحقيقات عادةً إلى مساعدة الوكالات الإخبارية والحكومات وجهات التحقُّق من الوقائع وشركات التواصل عبر الإنترنت وغيرها على فهم هذه الحركيّة التي لا تهدأ، واتخاذ التدابير المضادة الملائمة لمكافحتها. وهي تسعى إلى تنبيه الجهات الفاعلة إلى الابتكارات الجارية في مجال تكتيكات واستراتيجيات المعلومات المُضلّلة. ويجري تطوير منهجيات التصنيف بشكل متزايد للسماح بوضع فهارس أفضل تشمل الحوادث المتصلة ببعضها البعض[101].

1.2.4 من هي الأمور ومن هي الجهات المستهدفة؟

ترصد الاستجابات الاستقصائية طائفةً من الجهات الفاعلة. ويركز عدد كبير منها على حملات التّضليل الممولة أو المدعومة من الحكومات. وعلى الرغم من أن العديد من الاستجابات التجارية والصادرة عن الشركات قامت في البداية بالتركيز على دور حكومة أو حكومتين في حملات تضليل الهادفة، فقد توسعت لائحة الدول الراعية للتضليل (Nimmo, 2019; Francois et al., 2019; Nimmo et al., 2019a; Gleicher, 2019a;Gleicher, 2019b). ويحدد تقرير «الجرد العالمي للتلاعب المنظم بوسائل التواصل الاجتماعي لعام 2019» (Bradshaw & Howard, 2019) أو '2019 Global Inventory of Organised Manipulation Social Media حملات قامت من خلالها بعض الحكومات بالتلاعب بمواقع التواصل الاجتماعي لشن حملات ضد دول أخرى، مع الإشارة إلى أكثر من 26 دولة تعمد إلى القيام بأنشطة تضليل داخلية. غير أن الكيانات المذكورة أعلاه تميل إلى الإشاحة بنظرها عن الأنشطة السرية أو العلنية التي تقوم بها حكومات معينة في بلدان أجنبية أو محلياً. وهناك تركيز على المعلومات المُضلّلة التي تُروّج لها جهات غير رسمية مثل: الجماعات العنصرية التي تدعو إلى هيمنة العرق الأبيض في الولايات المتحدة أو الحركات اليمينية المتطرفة في الهند (the Southern Poverty Law Center, Equality Labs –see Soundararajan et al., 2019).

---

[101] https://www.bellingcat.com/

وينظر محققون آخرون إلى الجهات الفاعلة ذات الدوافع التجارية أو إلى الجهات الفاعلة المختلطة (وإن كانت تعمل في كثير من الأحيان لأغراض سياسية). على سبيل المثال، تُشارك «مزارع النقر» الأجنبية في حملات تضليل سياسية واسعة في الولايات المتحدة وذلك لأسباب تجارية (Silverman & Alexander, 2016, Soares, 2017). ومن الأمثلة الأخرى التحقيق في «مجموعة إبوك ميديا جروب» Epoch Media Group، وهي شركة تجارية لها دوافع سياسية، قادت حملة تضليل بما في ذلك عبر حسابات أو ملفات شخصية زائفة (Nimmo, et al., 2019b). وقد تم كشف هذه الحملة من خلال تقارير Snopes وعملية FFS# التي قامت بها غرافيكا/فيس بوك[102]. ومن الاتجاهات الهامة التي شهدتها السنتان إلى السنوات الثلاث الماضية نمو الجهات الخاصة المأجورة لنشر التّضليل والتي تقدم خدماتها للسياسيين وللأحزاب السياسية –كما هو موثق في عدد من البلدان (Ong & Cabañes, 2019; Silverman et al., 2020). وقد أظهر استطلاع أجرته Buzzfeed News في عام 2019 استناداً إلى عمليات إزالة الحسابات من قبل المنصات، والتحقيقات المنشورة من قبل شركات الأمن والبحوث «أنه منذ عام 2011،[103] تم إسناد 27 عملية معلومات عبر الإنترنت على الأقل جزئياً أو كلياً إلى شركات العلاقات العامة أو التسويق. ومن بين هذه العمليات، حدثت 19 منها في عام 2019 وحده». (Silverman, et al., 2020). كما أجرت صحيفة ديلي مافريك Daily Maverick في جنوب أفريقيا تحقيقاً مهماً آخر حول شركة العلاقات العامة Bell Pottinger التي تتخذ من المملكة المتحدة مقراً لها الآن، والتي تبيّن أنها تزرع بذور التوتر العنصري بشكل مصطنع في البلاد وسط فضيحة مرتبطة بالرئاسة تتعلق بالاستيلاء على الدولة للنهوض بالمصالح الخاصة (Thamm 2019; Posetti et al 2019a).

2.2.4 من هي الجهات التي تسعى الاستجابات الاستقصائية إلى مساعدتها؟

إنّ التقارير الصحفية الاستقصائية تخدم طائفةً من الجهات الفاعلة، بما في ذلك الشركات العاملة في مجال الكشف عن السلوكيات المُفتعلة المُعد لها على منصاتها والتحقيقات الرسمية. ومن أحد هذه التحقيقات التحقيق الذي أجراه الكونغرس الأمريكي في التدخل الأجنبي قبل وأثناء الانتخابات الأميركية في عام 2016 (U.S. Senate Select Committee on Intelligence, 2018)، واللجنة الدولية الكبرى المعنية بالتّضليل و«الأخبار الزائفة» التي أطلقتها المملكة المتحدة[104]. وتستخدم الحكومات أيضاً هذه التحقيقات، على سبيل المثال المقترحات السياسية للاتحاد الأوروبي المبنية على التقارير التي تتم بتكليف منه (EU Disinfo Lab, 2019b). وكثيراً ما ينعكس التنسيق بين شركات التواصل عبر الإنترنت والجهات الفاعلة الخارجية التي تحقق في حملات معينة من خلال علاقات التمويل بينها (يُرجَى مراجعة القسم أدناه والقسمين 4.1 و7.1).

كما وتضطلع منظمات المجتمع المدني، مثل منظمة العفو الدولية، بأعمال استقصائية جنائية تهدف إلى رصد المعلومات المُضلّلة المرتبطة بانتهاكات حقوق الإنسان وكشفها وصدّها. وتشمل هذه المؤسسات هيئة التحقُّق الرقمي التابعة لمنظمة العفو الدولية[105] – وهي شراكة مع ست جامعات دولية تتعاون

....

[102] https://graphika.com/

[103] https://www.militaryfactory.com/dictionary/military-terms-defined.asp?term_id=2637

[104] https://www.oireachtas.ie/en/press-centre/press-releases/20191025-international-grandcommittee-on-disinformation-and-fake-news-dublin-ireland-wednesday-6th-and-thursday-7thnovember-2019/

[105] https://www.theengineroom.org/digital-verification-corps/

أيضاً في التحقيقات الصحفية مفتوحة المصدر (Fortune, 2018). ومن بين أصحاب المصلحة الآخرين المواطنون وعدد متزايد من المشاركين في المجتمع العالمي لمُكافحة التّضليل. ومن الأمثلة على ذلك[106] مشروع «مُفكّكي الرُّموز» أو Decoders التابع لمنظمة العفو الدولية، الذي يحشد المساعدة التطوعية من الجمهور للمساعدة في التحقُّق من الادعاءات المتعلقة بانتهاكات حقوق الإنسان. كذلك، قامت «منظمة آفاز» العاملة في تنظيم حملات المناصرة بإجراء تحقيقات، بما في ذلك حول استجابات شركات التواصل عبر الإنترنت، فدعت إلى إجراء تغييرات وفقاً لذلك[107].

والدافع وراء الاستجابات الاستقصائية هو تحسين فهم أفعال التّضليل والحملات المحددة للسماح باتخاذ إجراءات لمكافحتها. ويمكن أن تشمل هذه الإجراءات عمليات إزالة المُحتوى أو خفض أهميته وإجراءات قانونية وتدابير تعزيز الشفافية والمُساءلة وإصلاح السياسات الناظمة أو سياسة الشركات. وتهدف الاستجابات الاستقصائية أيضاً إلى كشف الأساليب المعتمدة في حملات التّضليل لمنع استخدام هذه الأساليب من قِبل جهات فاعلة مأجورة، وضمان معرفة الباحثين في مجال المعلومات المُضلّلة بها. وهي تُحدّد التحديات البنيوية التي تطرحها المعلومات المُضلّلة بدلاً من البحث في أمثلة عن أعراض هذه الظاهرة أو عناصر المُحتوى الفردية. وكمثال على ذلك، تؤكّد مبادرة مختبر الاتحاد الأوروبي للمعلومات المُضلّلة EU Disinfo Lab التزامها «بالاستمرار في رصد أنشطة التّضليل على المنصات الرئيسية (الرقمية والتقليدية)، وتحديد الاتجاهات والتهديدات، وتنبيه الناشطين والباحثين منها، وإتاحة كافة الموارد على نطاق واسع للجهات الشريكة»[108].

وفي هذا الإطار الواسع، تتباين علاقة المنظمات التي تصدر تقارير التحقيق بالدولة. فهناك تعارض بين جهة متخصصة في نشر الأخبار على غرار Bellingcat[109] التي تركز على التحقيقات التي تستخدم البيانات المتاحة للجمهور والمصادر المفتوحة لإثبات الوقائع، ومنظمات التحقيق التي تعمل بشكل وثيق مع بعض الدول لتحقيق جدول أعمالها و/أو التي تعتبر تابعةً لشركات معينة.

3.2.4 ما هي النتائج التي تنشرها الاستجابات الاستقصائية؟

تتخذ التقارير الاستقصائية أشكالاً مختلفة. وتركّز معظم التحقيقات التي تجريها المنظمات غير الحكومية والمنظمات الصحفية على توفير معلومات عن آليات ومصادر المعلومات المُضلّلة (لكشف النهج المتبعة في اختلاق ونشر المعلومات المُضلّلة) ووضع تقارير متعمقة للاستخدام واسع النطاق. وتعد الشفافية في الأساليب المستخدمة (وفي الواقع المشاركة في وضع هذه التقارير عبر حشد الجمهور في عملية جمع الأدلة) جزءً أساسياً من مقاربات الاستخبارات من مصادر مفتوحة، المعروفة بمختصر OSINT، كما تمارسه BellingCat والمنظمات المختلطة الأخرى، في حين تنتج بعض المنظمات بيانات تستند إلى تخصصاتها (على سبيل المثال، أساليب تحليل شبكة الرسوم البيانية الاجتماعية في حالة Graphika). ولا توفر التحقيقات التي تجريها المجموعات الداخلية التابعة لشركات التواصل عبر

106 https://decoders.amnesty.org/

107 https://secure.avaaz.org/campaign/en/disinfo_hub/

108 https://www.disinfo.eu/

109 https://www.bellingcat.com/

الإنترنت بيانات كاملة عن الطريقة التي استخدمتها لتحديد المعلومات المُضلّلة. وإنّ سبب الانتقادات الموجهة ضد المقاربات التي تتبعها الشركات في تحديد المعلومات المُضلّلة ودعمها لشركات أخرى في هذا المجال هو أنها لا توفر بيانات تساعد على رصد حملات التّضليل وتصنيفها وتحديدها.

غير أن عدداً متزايداً من المجموعات يحاول وضع منهجيات مشتركة للتصنيف. ومن الأمثلة على ذلك العمل الذي قام به مختبر الأبحاث الرقمية الجنائية في مجال وضع إطار للتصنيف في مشروعه الخاص بالتّضليل المعنون «انقسامات المعلومات المُضلّلة» (Brooking, et al., 2020) المعروف باسم Dichotomies of Disinformation (بدعم من Jigsaw وهي وحدة في الشركة القابضة لغوغل المُسمّاة Alphabet). وتتضمن منهجية التصنيف أكثر من 150 متغيراً ثنائياً وكمّياً وقائماً على النص، مصنفاً تحت متغيرات أساسية هي: الهدف والمنصة والمُحتوى والطريقة والإسناد والقصد. وتشمل مبادرات التصنيف الأخرى المسوحات التي يجريها «مشروع الدعاية الحاسوبية» the Computational Propaganda Project حول التلاعب المنظم بوسائل التواصل الاجتماعي بناءً على تقييم التقارير الواردة في وسائل الإعلام الإخبارية (Bradshaw & Howard, 2019)، وتقارير «معهد المستقبل» أو Institute for the Future حول أنواع الاستفزاز على الإنترنت المكفول من قِبل الدولة في إطار حملات التّضليل والمضايقات عبر الإنترنت (Monaco & Nyst, 2018).

4.2.4 ما هي الجهات الفاعلة الرئيسية ومَن يُموّل الاستجابات الاستقصائية؟

هناك مجموعة من المبادرات التي تقوم بإجراء تحقيقات منظمة حول المعلومات المُضلّلة وإصدار تقارير متعمقة وهي تشمل:

الكيانات التي تركز بشكل أساسي على المعلومات المُضلّلة: يُعدّ مختبر الأبحاث الرقمية الجنائية التابع للمجلس الأطلسي مثالاً على منظمة تركز بشدة على تحديد حملات التّضليل وتحليل محتواها ومكافحتها. وهو ينشر التقارير، ويعمل على تطوير الخبرة الميدانية وتحديد المنهجيات ومقاربات التعقّب (Brooking, et al., 2020). في السياق الأوروبي، يعدّ مختبر الاتحاد الأوروبي للمعلومات المُضلّلة EU Disinfo Lab منظمةً غير حكومية حديثة الإنشاء؛ يُشغّل منصةً لتحليل المعلومات المُضلّلة في أوروبا، كما ويراقب أيضاً ويُحدّد حملات التّضليل ذات الأبعاد الدولية (EU Disinfo Lab, 2019a; Carmichael & Hussain, 2019). وتشمل الجهات الفاعلة التي تستجيب للتضليل في هذه الفئة الفرعية الجماعات التي تموّلها الحكومات والمنظمات غير الحكومية والبرامج الأكاديمية المتخصصة. وقد خاضت بعض التحقيقات في نماذج الأعمال المستخدمة من قبل العديد من شركات الإنترنت، فبيّنت أن هذه الشركات تميل إلى نشر الشائعات ونظريات المؤامرة. على سبيل المثال، صرّح مركز التكنولوجيا الإنسانية أو Center for Humane Technology أنّ يوتيوب قام باقتراح مقاطع فيديو من قِبل «أليكس جونز» أكثر من 15 مليار مرة وهي مقاطع تقوم بنشر نظريات المؤامرة.[110]

[110] http://humanetech.com/wp-content/uploads/2019/07/CHT-Undivided-Attention-Podcast-Ep.4-Down-the-Rabbit-Hole.pdf; https://www.newamerica.org/oti/events/online-getting-to-the-sourceof-the-2020-infodemic-its-the-business-model/

الكيانات التي تعتمد منهجيات ذات صلة بالتّضليل، مثل الاستخبارات مفتوحة المصدر (OSINT): بالتوازي مع تطوُّر البحوث والتحقيقات حول التّضليل في السنوات الست الماضية، شهد التحقيق عبر المصادر المفتوحة تطوراً كبيراً أي استخدام «المصادر المفتوحة» ووسائل التواصل الاجتماعي لإجراء تحقيقات في موضوعات مثل: جرائم الحرب واستخدام الأسلحة الكيميائية. ومن المنظمات التي تعمل في هذا المجال منظمة Bellingcat، التي تصف نفسها بأنها « ... مجموعة دولية مستقلة من الباحثين والمُحقّقين والمواطنين الصحفيين تستخدم المصادر المفتوحة ووسائل التواصل الاجتماعي للتحقيق في مجموعة متنوعة من الموضوعات من (...) مُرتزقة المخدرات والجرائم ضد الإنسانية إلى تعقب استخدام الأسلحة الكيميائية والنزاعات والصّراعات في جميع أنحاء العالم»[111]. وللمنظمة موظفون ومساهمون في أكثر من 20 دولة حول العالم، يعملون في مجال تتقاطع فيه التكنولوجيا المتقدمة مع الأبحاث الجنائية والصحافة والتحقيقات والشفافية ومُراقبة المُساءلة. وتشمل الكيانات في هذه المجموعة جماعات تموّلها مؤسسات وحكومات ومنظمات غير حكومية.

التحقيقات التي تجريها جهات رصد غير حكومية مع التركيز على حُريّة التّعبير المواضيعية أو القطاعية: على الرغم من أنه لا ينبغي الخلط بين المعلومات المُضلّلة وخطاب الكراهية، فهناك مجموعات تعمل على المسألتين وتحقق في أنماط مشاركة المعلومات المُغرِضة في سياقات مواضيعية خاصة. ومن الأمثلة على ذلك المركز القانوني الجنوبي للفقر أو Southern Poverty Law Center في الولايات المتحدة، الذي أُنشئ «لرصد جماعات الكراهية والجماعات المتطرفة الأخرى في جميع أنحاء الولايات المتحدة، ولكشف أنشطتها أمام الجمهور ووسائل الإعلام وسُلطات إنفاذ القانون»[112]. ويصدر المركز تقريراً شاملاً نصف سنوي عن وضع هذه الحركات وأنشطتها، كما ويصدر تقارير محددةً عن أنشطة دعائية معينة. كذلك[113]، يوثق «معهد الحوار الاستراتيجي» Institute for Strategic Dialogue الذي يتخذ من «لندن» مقراً له تقارير عن العنف المتطرف وما يتصل به من خطابات. (كما يبيّن القسم 7.1 من الفصل السابع، هناك تداخلات كبيرة بين الاستجابات المعيارية والأخلاقية للتضليل، ومسألة خطاب الكراهية).

تحقيقات متعمقة تجريها وسائل الإعلام الإخبارية: تقوم مجموعة من وسائل الإعلام الإخبارية بتحقيقات مستمرة في المعلومات المُضلّلة. ومن الأمثلة على ذلك Buzzfeed News، التي تقدم رؤى وتحقيقات حول حملات التّضليل الفردية والاتجاهات المتبعة في المعلومات المُضلّلة، مثل: الاستخدام المتزايد لشركات العلاقات العامة المأجورة لنشر المعلومات المُضلّلة (Silverman, et al., 2020).

وقد أجرت وسائل إعلام إخبارية أخرى تحقيقات متعمقة حول حملات معينة، على غرار Rappler التي قامت بمسح شبكات المعلومات المُضلّلة أثناء الانتخابات الرئاسية عام 2016 في الفلبين وبعدها (Ressa 2016; Posetti et al., 2019a)، وتحقيقات صحيفة Daily Maverick في جنوب أفريقيا المشار إليها أعلاه (Thamm 2019; Posetti et al., 2019a)، وتحقيقات الشبكة الأفريقية لمراكز التحقيق الاستقصائي أو African Network of Centers for Investigative Reporting حول التلاعب بوسائل الإعلام في جنوب

111 https://www.bellingcat.com/about/

112 https://www.splcenter.org/fighting-hate

113 https://www.isdglobal.org/isdapproach/

أفريقيا[114] (ANCIR, n/d). كذلك، يقوم موقع CodaStory بالتحقيقات وهو موقع متخصص في المعلومات المُضلّلة يركّز على التحقيقات في الحملات المنظمة المرتبطة بالجهات الفاعلة الحكومية ووكلاء التّضليل المأجورين (Dorroh 2020).[115]

البحث الأكاديمي الموجّه نحو العمل: ينتج عدد متزايد من الأقسام الأكاديمية تحليلات وصفية للاستراتيجيات المتبعة في حملات التّضليل، مثلاً جرد حملات التلاعب المنظمة على وسائل التواصل الاجتماعي (استناداً إلى تحليل مُحتوى وسائل الإعلام الإخبارية) الذي أصدره مشروع الدعاية الحاسوبية أو Computational Propaganda Project في معهد أكسفورد للإنترنت (Bradshaw & Howard, 2019). كذلك، تصدر أبحاثاً مفصلةً حول استراتيجيات محددة وسياقات قُطريّة. ومن الأمثلة على هذه الأبحاث بحث أكاديمي حول التلاعب المنظم على وسائل التواصل الاجتماعية في الفلبين (Ong & Cabañes, 2018). ويركز بحث آخر من «معهد رويترز لدراسة الصحافة» بجامعة أكسفورد على تعرض الجمهور للتضليل في عدد من البلدان، وكذلك على الاستجابات الاستقصائية الصحفية (Posetti et al., 2019a) وأثار وسائل الإعلام (Nelson & Graves, 2017).

الكيانات التجارية التي تعمل في مجال تحليل الشبكات الاجتماعية والأمن السيبراني: تقدّم مجموعة من الشركات التجارية خدمات أو تجري بحوثاً استقصائيةً في حملات التّضليل. ومن الأمثلة على ذلك Graphika، التي تركّز على كشف «حملات التأثير الاستراتيجية التي تتم عبر الإنترنت وعلى نطاق واسع وذلك من خلال تحليل الاضطرابات التي تظهر في الشبكة وتحديد الأمور التي تنتشر على خرائط الشبكة بدرجة عالية من الانتقال التي قد تصل بسرعة إلى الانتشار الواسع».[116] وتطبّق الشركة تحليل شبكات التواصل الاجتماعي لإجراء تحقيقات في حملات مُحددة. وتتم هذه التحقيقات أحياناً بالتنسيق مع جهات فاعلة أُخرى، على سبيل المثال (Nimmo et al., 2019b)، مع شركات مثل فيس بوك، في عملية «#FFS: Fake Face Swarm» وهي تحليل لملفات/حسابات شخصية مرتبطة بمجموعة «إبوك ميديا». ومن الأمثلة الأخرى على الكيانات التجارية شركة FireEye، التي لديها خلفية في الأمن السيبراني التجاري. وقد حددت وحققت في انتهاكات الأمن السيبراني وحملات التّضليل المتصلة بها في دول مختلفة (Revelli & Foster, 2020).

التحقيقات التي يُجريها فريق الحد من المخاطر داخل الشركات: لجميع شركات وسائل التواصل الاجتماعي الرئيسية فرق تعمل على تحليل التهديدات، وفرق مخصصة للحفاظ على «سلامة الموقع» أو تحديد «السلوكيات المُفتعلة المُعد لها» (Gleicher, 2018a). فعلى سبيل المثال، أصدر فيس بوك تقريراً حول معالجة السلوك المُفتعلة المُعد له في عدد من البلدان (Gleicher, 2020). وتشارك هذه الفرق أحياناً بيانات محددة مع شركاء خارجيين أو تتعاون/تتعاقد مع شركات خارجية ومجموعات غير ربحية/أكاديمية (Gleicher, 2018b). وفي حالة فيس بوك، يتم التعاون مع أنواع أخرى من الكيانات المذكورة في هذا الفصل، على سبيل المثال: مختبر الأبحاث الرقمية الجنائية، وGraphika، وFireEye.

---

114 https://s3-eu-west-1.amazonaws.com/s3.sourceafrica.net/documents/118115/Manufacturing-Divides.pdf

115 يُرجَى مراجعة المناقشة المصوّرة بالفيديو حول التحقيقات الصحفية المتعمقة في المعلومات المُضلّلة في سياق جائحة فيروس كورونا التي تجمع محررة CodAStory «ناتاليا أنتيليفا»، والاختصاصية في المعلومات المُضلّلة في Buzzfeed «جاين ليتفيننكو»، والمحررة التنفيذية في Rappler «ماريا ريسا»: https://www.youtube.com/watch?v=tBp4OKSW_ho&feature=youtu.be

116 https://www.graphika.com/graphika-labs

وكما يتبين من هذه الأمثلة، تموّل مجموعة من الجهات الفاعلة التقارير الاستقصائية المتعلقة بالتّضليل. وتتلقى الجهات الفاعلة غير الربحية وغير الحكومية تمويلاً من مؤسسات ومن شركات ومن حكومات أيضاً. وتتمتع بعض الجهات الفاعلة بوضع مؤسَّسي أكثر في هذا الصدد. فـ«مختبر الأبحاث الرقمية الجنائية في المجلس الأطلسي» على سبيل المثال، هو جزء من كيان أكبر يتلقى تمويلاً كبيراً من حكومات المملكة المتحدة والولايات المتحدة الأميركية والإمارات العربية المتحدة. كما ويحصل على دعم إضافي من فيس بوك (Lapowsky, 2018). كما وتحظى كيانات أخرى مثل: Bellingcat على التمويل من مؤسسات معيّنة وهي تركّن إلى التدريبات وورش العمل التي تقدمها كمصدر دخل مكمّل لدخلها الأساسي. وتشارك بعض الوكالات الإخبارية القديمة أيضاً في التحقيقات التعاونية حول المعلومات المُضلّلة ممّا من شأنه أن يجتذب التمويل من الجهات المانحة (على سبيل المثال: من خلال الاتحاد الدولي للصحافيين الاستقصائيين (International Consortium of Investigative Journalists)، في حين تجري منظمات أخرى تحقيقات مستقلة تندرج في إطار مهمة الصحافة، وهي محاسبة السلطة.

ومن شأن الاستجابات الاستقصائية التعاونية ومتعددة التخصُّصات التي تجمع مثلاً بين الجهات الفاعلة التي تتمتع بالخبرة في المجالات المذكورة أعلاه، أن تزيد من فعالية هذه التدخلات. فعلى سبيل المثال، شملت التحقيقات الصحفية التي أجراها موقع Rappler في الفلبين شراكات مع منظمات غير حكومية وأكاديميين وخبراء في التكنولوجيا.

5.2.4 دراسة حالة بشأن الاستجابة: المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا

نظراً لطبيعة الاستجابات الاستقصائية حول جائحة فيروس كورونا التي لابد أن تكون أكثر تعمقاً وتتطلّب الكثير من الموارد، ونظراً لحداثة الوباء، كان هناك عدد أقل من الاستجابات الاستقصائية المنشورة حول جائحة فيروس كورونا في أيار/مايو 2020 مقارنةً بجهود التحقُّق من الوقائع والتدقيق. ومع ذلك، شملت المنظمات المتخصصة في الاستجابات الاستقصائية مخرجات من العديد من المنظمات غير الحكومية[117] والجهات الناشرة للأخبار (Evans, 2020)، ومراكز الفكر (EUvsDisinfo, 2020)، وتحقيقات مشتركة بين الأكاديميين ووسائل الإعلام المستقلة (Hollowood & Mostrous, 2020). وتشمل الموضوعات التي تم التحقيق فيها حملات التّضليل المتعلقة بجائحة فيروس كورونا التي أطلقتها وسائل إعلام مدعومة من الحكومات، ومن الحركات المتطرفة العنيفة، ومن المجموعات المناهضة للمهاجرين، ومن الشبكات اليمينية المتطرفة (ISD, 2020a). وتعمل هذه المنظمات عبر شركات التواصل الاجتماعي الرئيسية، بما في ذلك تويتر (منشورات مفتوحة ورسائل مباشرة)، وفيس بوك (حسابات، ومجموعات، وصفحات، وتطبيق ماسنجر)، ويوتيوب (مقاطع فيديو وتعليقات)، وواتساب، وانستغرام (منشورات مفتوحة ورسائل خاصة)، على الرغم من الجهود التي تبذلها هذه الشركات لمواجهة أزمة المعلومات المُضلّلة.

ولا يضم معظم التحاليل حتى الآن تحقيقات معمّقة من قبل مؤسسات أو مراكز فكر أو كيانات تجارية، بل تقارير من وسائل إعلام إخبارية، على سبيل المثال ProPublica (Kao, 2020) وصحيفة نيويورك تايمز (New York Times, 2020) في الولايات المتحدة، وRappler في الفلبين (Gaw, 2020). وقد وردت تقارير

117 https://rsf.org/en/disinformation

كثيرة في وسائل الإعلام حول الانتشار العلني والمستمر للمعلومات المُضلّلة من قبل القادة السياسيين خلال الجائحة، إلى جانب تقييمات حول كيفية استخدام الإحصاءات لخلق انطباعات مضللة.

وتشمل فئة أخرى من الاستجابات الاستقصائية للتضليل حول جائحة فيروس كورونا إرشادات حول أنواع المعلومات المُضلّلة التي تم تحديدها حتى الآن، على غِرار موجزين سياسيين حول أزمة المعلومات المُضلّلة؛ نشرتهما اليونسكو بالشراكة مع المركز الدولي للصحافيين (Posetti & Bontcheva, 2020). وقد حدد هذان الإيجازان تسعة أنواع من المعلومات المُضلّلة حول جائحة فيروس كورونا، وأربعة عوامل ناقلة رئيسية، وعشر طرق استجابة. الرجاء مُراجعة الأبحاث التي أجراها «معهد رويترز» حول أنواع المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا ومصادرها والادعاءات المتعلقة بها (Brennen et al., 2020) والتي حددت القادة السياسيين والمشاهير كمصادر رئيسية للتضليل.

كذلك، كانت فرق للحد من المخاطر داخل شركات التواصل عبر الإنترنت، تقوم بتحقيقات في المعلومات المُضلّلة المتعلّقة بجائحة فيروس كورونا إمّا بشكل مستقل، أو بالتعاون مع جهات خبيرة أخرى (Shu & Shieber, 2020). ولكنّ النتائج التي تم الكشف عنها جاءت مجزأةً،[118] ووجد الصحفيون المتخصّصون أنها ناقصة (Turvill, 2020).

6.2.4 كيف يتم تقييم الاستجابات الاستقصائية؟

تقوم جهات فاعلة كثيرة بالإفصاح بكل شفافية عن الأساليب والإجراءات التي تتبعها، وتنشر تقارير متاحة للجمهور عن النتائج التي توصّلت إليها. غير أنّ التقييمات الصريحة للأثر والفعالية لا توضع في متناول الجمهور من قِبل معظم الجهات الفاعلة المشاركة في إعداد التقارير الاستقصائية حول المعلومات المُضلّلة. ومن نتائج التحقيقات التي تظهر للجمهور نتائج التحقيقات التي يجريها قطاع ما والتحقيقات التعاونية حول حملات التّضليل؛ حيث تحدث عمليات إزالة الحسابات والمُحتوى ذات الصلة بالتحقيق من على فيس بوك أو تويتر أو من على منصات أخرى للتواصل الاجتماعي. كذلك، في سياق التحقيقات التي تتم بتكليف من الحكومة، على سبيل المثال في ما يخص التدخل الأجنبي في انتخابات الولايات المتحدة عام 2016؛ يتم استخدام البيانات مباشرةً في جلسات الاستماع في الكونغرس.

7.2.4 التحديات والفرص

تتجاوز التقارير الاستقصائية عمليات منفردة لفحص الوقائع وكشف زيف المعلومات حيث أنها تسعى إلى صياغة رؤى وتحاليل أعمق وتفاصيل عن حملات مُحددة. ومع نضوج هذا المجال، تطوّرت القدرة على تتبع الجهات الناشرة للمعلومات المُضلّلة على مر الزمن. ويمكن في هذا المجال النظر في الابتكارات أو المقاربات التي استُخدمت في التدخل الأجنبي في الانتخابات الأميركية عام 2016 (U.S. Senate Select Committee on Intelligence, 2018) من خلال حملات على غرار IRA CopyPasta كما ذُكر في (François, et al., 2019).

....

118 https://en.unesco.org/sites/default/files/unesco_covid_brief_en.pdf

ويتمثل أحد التحديات في أن الصحفيين الذين يُجرون تحقيقات في المعلومات المُضلّلة يتعرّضون لهجمات تستهدفهم، كمضايقات إلكترونية ومعلومات مضللة موجهة ضدهم (Posetti, 2017)، كما في حال «ماريا ريسا» العاملة في Rappler. وقد قدم عدد من شركات التواصل عبر الإنترنت بعض الدعم في هذا المجال على غِرار نصائح السلامة التي أصدرتها لجنة حماية الصحفيين التابعة لفيس بوك حول حماية المصادر[119] وجهات الاتصال، ومشروع Shield من غوغل[120]. ومع ذلك، كان هناك انتقاد لتأخر الشركات في الرد على الشكاوى المتعلقة بالمضايقات، وجعل الضحية مسؤولةً عن حماية نفسها وعن حظر الجهات المُضايقة والإبلاغ عنها وحذفها بدلاً من أن تتخذ الشركة إجراءات سريعة في هذا الخصوص (Posetti, 2020).

ومع انتقال المنظمات إلى دراسة طبيعة حملات المعلومات المُضلّلة من خلال وضع رموز وتقدير كمي لها، بدأت تظهر مجموعة من البيانات تسمح بإجراء تحليل مقارن (كما هو مذكور أعلاه). كما ويشارك المزيد من المنظمات اليوم في نشاطات تثقيف الجمهور، إلى جانب وضع التقارير والتحقيقات المكثفة، وذلك لضمان تبادل الممارسات المُميّزة والمقاربات الجديدة الهادفة إلى التصدي للمعلومات المُضلّلة. ومن الأمثلة على ذلك «المؤتمر السنوي للمحققين الرقميين لتبادل أساليب العمل Digital Sherlocks methods-sharing conference الذي ينظمه مختبر الأبحاث الرقمية الجنائية[121]، و«المؤتمر السنوي لمختبر الاتحاد الأوروبي للمعلومات المُضلّلة»، وأساليب التحقيق من المصادر المفتوحة والتدريبات التي تجريها منظمة Bellingcat.

ومع ذلك، تواجه التحقيقات المتعمقة تحديات كبيرة تتجاوز مسألة التكلفة والتعقيد. فمعظم التحقيقات يتم من دون النفاذ إلى مجموعات البيانات الكاملة اللازمة لفهم حملة معينة بشكل كامل؛ لأن شركات التواصل عبر الإنترنت لا تُفصح عن هذه البيانات بشكل روتيني. وقد شرح تويتر سياسته الخاصة بالإفصاح عن البيانات في مقال صادر عن رئيس وحدة سلامة الموقع (Roth, 2019). وقد انتقد الباحثون فيس بوك لتأخره في توفير النفاذ إلى البيانات ولكنه أصدر مؤخراً مجموعة بيانات متماشية أكثر مع التزاماته (King & Persily, 2020). ومن المسائل الأخرى عدم السماح بالنفاذ إلا لعدد محدود من الباحثين، وهم غالباً ممن يتلقون منحاً كبيرةً من هذه الشركات.

ولا تتوفر للباحثين إلا معلومات وأدوات محدودة مما لا يسمح لهم القيام بتحليل عابر للمنصات، على الرغم من أنّ عدداً قليلاً من حملات التّضليل (أو المعلومات المُضلّلة واسعة الانتشار) يقتصر على منصة واحدة. وهناك مشكلة خاصة تتمثل بمسألة النفاذ إلى البيانات حول المعلومات التي يتم تبادلها على تطبيقات المراسلة – التي تنتشر عليها المعلومات المُضلّلة – والتي غالباً ما تكون مُشفّرةً من طرف إلى طرف لأسباب تتعلق بالأمن والخصوصية. ومع ذلك، فإن هذه الشركات لديها إمكانية النفاذ إلى البيانات الوصفية حول الحركة والمجموعات على التطبيق، حتى لو لم يكن لديها إمكانية الوصول إلى رسائل محددة. ويمكن أن يساعد النفاذ إلى هذه المعلومات المحققين في الكشف عن الأنماط المستخدمة في كل شبكة من شبكات المعلومات المُضلّلة.

119 https://www.facebookblueprint.com/student/path/188883-journalist-safety

120 https://blog.google/outreach-initiatives/google-news-initiative/our-efforts-help-protect-journalistsonline/

121 https://digitalsherlocks.org

وفي السنوات الأربع الماضية، كان هناك تركيز أوّلي شديد على المعلومات المُضلّلة المنسوبة إلى دولة واحدة على وجه الخصوص. ومع ذلك، وكما هو مبيّن أعلاه، بدأت استجابات الشركات والاستجابات الأكاديمية والاستقصائية بالتركيز على طائفة أوسع من الدول والجهات الفاعلة الخاصة/الحكومية.

وكذلك، وكما ذُكر أعلاه، هُناك ثغرات كبيرة في النفاذ إلى المعلومات لدعم المجتمع المدني والصحافة والأوساط الأكاديمية على نحوٍ ملائم في مجال فهم حملات التّضليل العابرة للمنصات، وحملات التّضليل القائمة على تطبيقات المراسلة. وهذا يعطي شركات التواصل عبر الإنترنت فرصة التعاون مع الباحثين ومنظمات المجتمع المدني ذوي المهارات المتخصصة في هذا المجال لتحليل البيانات وتطوير السياسات.

8.2.4 توصيات بشأن الاستجابات الاستقصائية

يسمح هذا القسم باستخلاص عدد من التوصيات حول الإجراءات الواجب اتخاذها من قبل مجموعة من الجهات الفاعلة. وهي تشمل التالي:

- لا بد لجميع أصحاب المصلحة أن يُدركوا الحاجة إلى الاستثمار في الصحافة الاستقصائية النقدية والمستقلة كتدبير دفاعي ضد التّضليل، خاصّةً وأن الضغوط المالية التي وضعتها جائحة فيروس كورونا وجّهت ضربات قاضية إلى وسائل الإعلام الإخبارية في جميع أنحاء العالم، وهي تُهدّد مبادرات الصحافة الاستقصائية التي تعد مكلفةً.

- على شركات التواصل عبر الإنترنت أن تُتيح فرصاً أوسع للنفاذ إلى مجموعات بياناتها للباحثين المستقلين الذين يدرسون المعلومات المُضلّلة، بمن فيهم أولئك الذين لا يحصلون على تمويل بحثي كبير من هذه الشركات، وذلك لصالح تبادل المعارف لمُكافحة المعلومات المُضلّلة.

- **توصيات للمانحين ومنظمات البحوث:** زيادة الاستثمار في التحقيقات التعاونيّة ومتعددة التخصُّصات، وتعزيز التعاون بين الباحثين الأكاديميين وعلماء البيانات التجارية والمنظمات غير الحكومية والوكالات الإخبارية.

- تمويل التحقيقات السريعة في المعلومات المُضلّلة خلال حالات الطوارئ، مثل جائحة فيروس كورونا.

5

الفصل الخامس: الاستجابات الخاصة ببيئة المعلومات المُضلِّلة والتي تستهدف مُنتجي المعلومات المُضلِّلة وناشريها



1.5 الاستجابات التشريعية، وما قبل التشريعية والسياستية

المؤلفون: تريشا ماير، كلارا هانو، جولي بوزيتي، دينيس تيسو

يناقش هذا القسم الاستجابات التشريعية وما قبل التشريعية والسياستية التي تصدر عن الجهات الحكومية (التشريعية والتنفيذية والقضائية) والتي تشمل التدخلات التنظيمية لمواجهة التّضليل. وتغطي هذه الاستجابات أنواع مختلفة من الإجراءات التنظيمية، بدءً من التحقيقات واقتراحات القوانين وانتهاءً بالتشريعات وإنفاذ القانون. وهي عادةً تهدف إلى استغلال الدولة لسلطتها لرسم شكل البيئة المحيطة بإنتاج ونقل واستقبال المُحتوى، مما يؤثر إما على الدورة بأكملها، أو على مواقع وجهات معينة داخل هذه الدورة.

ويتم التصدي للمعلومات المُضلّلة على الإنترنت من منظورات متعددة بما في ذلك من خلال التشريعات القائمة التي لا تتناول التّضليل بشكل حصري ولكنها تقوم بمعالجة بعض جوانب هذه الظاهرة. ولا يمكن لهذا القسم أن يغطيها بشكل شامل، ولكن من المهم تسليط الضوء على بعض الوسائل المستخدمة لفهم السياق القانوني والسياسي الأوسع الذي تتبلور فيه الاستجابات الحكومية للتضليل. ولذلك يتم التركيز هنا على التشريعات والسياسات المرتبطة حصراً بالمعلومات المُضلّلة، إلا إذا كان هناك تشريع أو إجراء سياسي معيّن قد تمّ توسيع إطاره أو إعادة توجيهه ليغطي مسألة المعلومات المُضلّلة.

ولا شك في أنّ المقاربات المؤسسية والفردية المبنية على التنظيم الذاتي هي الطريقة الأساسية للتصدي للمعلومات المُضلّلة، إلا أنّ بعض الجهات الحكومية ترى أنه من الضروري بلورة تدخلات تنظيمية أيضاً. وقد تكون بعض هذه المقاربات مقيّدةً، في حين أن المقاربات الأُخرى قد تكون محفزة (ولكن في أحيان أقل). والقصد من ذلك هو توفير عوامل ردع وعوامل تحفيز (في أحيان أقل) تهدف إلى تغيير سلوك الجهات الفاعلة. ويتأثر شكل هذه الاستجابات بالتقاليد القانونية الوطنية/الإقليمية، وقوة الأُطُر القانونية والمعيارية الدولية، والحساسيات الثقافية.

من ناحية البُعد القسري لهذا النوع من التدخّلات، لابد من الإشارة إلى أنّ القوانين المطبقة على التّضليل غالباً ما تكون عامة، ممّا يهدد بتقييد التّعبير المشروع وإخضاعه للرقابة، بما في ذلك العمل الصحفي. كذلك، لابدّ من النظر في القواعد التي تتصدّى للتعبير الضار (على سبيل المثال، الادعاءات الكاذبة التي تهدف إلى بيع المنتجات) لمعرفة ما إذا كانت كافيةً أو ما إذا كانت هناك حاجة إلى وضع أنظمة جديدة، وكيف يمكن تجنُّب تقويض حُريّة التّعبير المشروع. وبالتزامن مع ذلك، لابد من البحث في ما إذا كانت هناك أنظمة قانونية فعّالة تحظر التحريض على التهجم على حُريّة الصحافة وسلامتها (بما يتضمّن مُروّجي التّضليل).

أما في ما يتعلق ببعض التدخلات التنظيمية التي تركّز لا على المنع بل على التحفيز، لابد من البحث إلى أي مدى تعتمد الشفافية والمساواة كمبدأ قانوني أساسي. ومثال على ذلك هو ما إذا كانت هناك أنظمة مفتوحة وعادلة تضمن تخصيص الأموال العامة لعمليات التحقُّق من الوقائع والخطاب المعاكس (مراجعة القسم 5.2 أدناه)، أو لوسائل الإعلام، وما إذا كانت هذه الأنظمة تضمن عدم إساءة استخدام هذه الأموال لأغراض سياسية.

المنهجية والنطاق

بهدف تحديد الاستجابات التشريعية وما قبل التشريعية والسياستية المتعلقة بالمعلومات المُضلّلة؛ استخدم هذا التقرير البحثي ثلاث مراجع تغطي مجموعةً واسعةً من البلدان والمقاربات كنقطة انطلاق، وهي: «دليل إجراءات مُكافحة التّضليل حول العالم» (من بوينتر، ويتم تحديثه بانتظام[122])، وتقرير مكتبة الكونغرس حول «مبادرات مُكافحة الأخبار الزائفة في بلدان مختارة» ([123]Library of Congress, 2019) وتقرير جامعة أكسفورد حول «مبادرات مُكافحة التّضليل» (Robinson et al., 2019)

وفي تحليل هذه الاستجابات التنظيمية، قام الباحثون بالرجوع إلى المصادر الأولية (القوانين، ووثائق السياسات، والبيانات الصحفية الحكومية، والمواقع، وغيرها) لفهم المبادرات الحكومية قدر الإمكان. ومتى تعذّر العثور على المصادر الأولية أو متى كان من الضروري الحصول على معلومات إضافية، تمت مراجعة مصادر ثانوية (مقالات إخبارية وتقارير أكاديمية وتحليلات قانونية.. إلخ). ولاعتبار المعلومات المُستقاة من مصادر ثانوية ذات موثوقية؛ اشترط الباحثون أن تكون قد نُشرت على مواقع متعددة. وقد سمحت هذه المصادر الثانوية بتحديد استجابات حكومية إضافية.

وهنا، لابد من الإشارة إلى أنّ بعض الدول قامت باقتراح أو سنّ تشريعات خاصة بالتّضليل. فيما قامت دول أخرى بإجراء تعديلات على قوانين أخرى غير مُخصَّصة للتضليل، حيث اعتبرت أنّ الأساس القانوني لمعالجة التّضليل موجود في قوانين قائمة، مثل: القانون الجنائي، والقانون المدني، والقانون الانتخابي أو قانون الأمن السيبراني. ومن المُسلّم به أن هناك أحكاماً تتعلق بالتّضليل والمعلومات المغلوطة و«الأخبار الزائفة» والأكاذيب والشائعات وما إلى ذلك في تشريعات كثيرة لا يمكن تغطيتها في تقرير واحد. وقد شمل التقرير الحالات التي تم فيها مؤخراً اقتراح أو اعتماد أو تطبيق تشريعات أو تعديلات على تشريعات تتعلق بالتّضليل. كذلك شمل الحالات التي ذُكر التّضليل فيها بوضوح في التقارير والحوارات والنقاشات التي أدت إلى اقتراح القانون أو إلى اعتماده أو تطبيقه. وفي بعض الحالات اعتمدت الدول «تدابير إيجابية» أي غير عقابية سوف تتم مناقشتها في القسم 5.2.

1.1.5 ما هي الأمور ومن هم الأشخاص الذين ترصدهم وتستهدفهم الاستجابات التشريعية؟

لفهم التحديات التي يطرحها استخدام الاستجابات التشريعية والسياستية على حُريّة التّعبير، لابد من الإشارة إلى الحق في حُريّة الرأي والتّعبير، على النحو المنصوص عليه في المادة 19 من الإعلان العالمي لحقوق الإنسان وفي العهد الدولي الخاص بالحقوق المدنية والسياسية:

> „ لكل شخص الحق في حُريّة الرأي والتّعبير، ويشمل هذا الحق حُريّة اعتناق الآراء من دون أي تدخّل، واستقاء الأخبار والأفكار وتلقيها وإذاعتها بأية وسيلة كانت دون تقيد بالحدود الجغرافية. "

....

[122] https://www.poynter.org/ifcn/anti-misinformation-actions/

[123] https://www.loc.gov/law/help/fake-news/index.php

لذلك، ينبغي تقييم التدابير التنظيمية التي تسعى إلى تقييد التّضليل بناءً على المعايير الدولية التي تؤكد على أن أي قيود تفرض على حُريّة التّعبير لابد أن يكون منصوص عليها بموجب قانون، وأن تثبت ضرورتها لأغراض مشروعة، وأن تكون الوسيلة الأقل تقييداً لتحقيق الهدف. كما وينبغي أن تكون محدودةً في الزمن إذا كانت تدابير اعتُمدت في حالات الطوارئ.

وللنظر في كيفية تأثير القوانين والسياسات في السياقات الإلكترونية على حُريّة التّعبير، يمكن تقييم الاستجابات التي تستهدف سلوكيات مختلف الجهات الفاعلة. وتسعى بعض الاستجابات إلى طرح تدابير قد تُفهم على أنها تدابير «إيجابية» -أي شروط لتطبيق الحق في حُريّة التّعبير. ولكنّ معظم التدابير تهدف إلى ردع الأشكال المسيئة لحُريّة التّعبير كما هي محددة في القانون، وبالتالي تُنتج ما يمكن تسميته بتدابير «سلبية».

ويتم اعتماد معظم هذه التدابير لحماية المواطنين. فمن جهة هناك خطوات على غرار قواعد حماية البيانات وسياسات تعزيز الدراية الإعلامية والمعلوماتية التي تسعى إلى تزويد الأفراد بمستوىً معيّن من الحماية والمهارات الأساسية للمشاركة في الحياة الإلكترونية. وفي الوقت نفسه، هناك قيود على التّعبير الذي يُسبّب الأذى للآخرين، مثل: التحريض على الكراهية والعنف (على أساس الانتماء العِرقي أو الإثني أو الجنس أو الدين)، والتشهير والدعاية النازية (في أُطُر محددة)، أو المضايقات والتهديد بالعنف. هذه القيود على حُريّة التّعبير مبررة وفقاً للمعايير الدولية، على الرغم من أن «مبادئ الرباط للمفوضية العليا لحقوق الإنسان» تُقيم بعض الفوارق البسيطة المهمة في هذا المجال بهدف وضع عتبة عالية لفرض هذه القيود[124]. وتستهدف هذه القيود ثلاثة أنواع من السلوكيات من بين أمور أخرى.

أولاً، يتم فرض تدابير عقابية على الأشخاص المتورطين في إنتاج وترويج وتوزيع المحتويات الضارة عند تجاوزهم القيود المفروضة على التّعبير. ولكن قد يؤثر ذلك بشكل غير مقصود على التّعبير الشرعي الذي، وإن كان مغلوطاً (كما في التهكم) أو مزعجاً (كما في الصدمة)، لا يُعتبر بالضرورة غير قانوني وفقاً للمعايير الدولية. كذلك لابد من معرفة ما إذا كانت هذه التدابير، من خلال تصميمها أو تطبيقها، قد وُضعت فعلاً لحماية الجمهور أو لحماية بعض المصالح الخاصة على غِرار مصالح شاغلي المناصب السياسية. بالإضافة إلى ذلك، من مضاعفات هذا النوع من القيود على حُريّة التّعبير أنّ هذه القيود تطبّق على المستوى الوطني وتتطلّب تعاون الشركات العالمية للتواصل عبر الإنترنت؛ والتي أصبحت من النواقل الرئيسية للمعلومات المُضلّلة واسعة الانتشار.

ثانياً، إنّ قواعد المنافسة وحماية المستهلك، مصحوبةً بالقواعد التي ترعى عمل القطاع، بما في ذلك القوانين المطبقة على الإعلانات المُضلّلة، هي التي ترسم معالم السلوك الاقتصادي المقبول لشركات التواصل عبر الإنترنت. لكن، كما هو وارد في القسم 6.3 حول الاستجابات الهادفة إلى حجب الربح المالي، تزداد التساؤلات في دوائر وضع السياسات عما إذا كانت القواعد الحالية تردع بما فيه الكفاية الاستغلال الاقتصادي للمحتويات المثيرة و/أو الكاذبة.

---

[124] خطة عمل الرباط بشأن حظر الدعوة إلى الكراهية القائمة على الانتماء القومي أو العنصري أو الديني التي تشكل تحريضاً على التمييز أو العداوة أو العنف: https://www.ohchr.org/EN/Issues/FreedomOpinion/Articles19-20/Pages/Index.aspx

ثالثاً، يخضع السلوك التقني للسياسات السيبرانية التي تتم صياغتها قانونياً والتي تسعى إلى ردع استخدام تقنيات الإنترنت لأغراض خبيثة، على غرار حملات المعلومات المُعد لها أو المتطفلة المستخدمة لأغراض التّضليل. كذلك، لابد من الإشارة إلى التعاون المتزايد في موضوعات مثل: مُكافحة الإرهاب، بهدف تبادل المعرفة والممارسات بين الجهات الفاعلة الحكومية والتقنية، ضمن الأُطُر القانونيّة حول الإرهاب.

رابعاً، إنّ التدخلات التنظيمية لتوجيه سلوكيات الفاعلين السياسيين تشمل قواعد ترعى الإعلانات الانتخابية وإعلانات الحملات السياسية.

ومن جهة التمكين، بدلاً من التدابير السياسية المقيّدة، قد يكون هناك تدخلات تنظيمية تهدف إلى تعزيز توافر المعلومات كبديل للتضليل. ويمكن أن تشمل هذه التدخلات تعزيز الشفافية والكشف الاستباقي من قبل المسؤولين، وتعزيز النفاذ إلى أنظمة المعلومات. ويمكن أن تشمل أيضاً تخصيص الأموال العامة لدعم وسائل الإعلام الإخبارية، ومبادرات التحقُّق من الوقائع، وحملات مُكافحة التّضليل من قبل كيانات خاصة أو عامة.

ولا شك في أنّ السياسات والاستجابات القانونية المتخذة على مستوى الدولة وبين الحكومات متقاطعة وتغطي جميع أنواع الأعمال. واستناداً إلى التحليل أعلاه، يمكن تحديد أربع مجموعات تستهدفها الاستجابات السياسية.

أولاً: إنّ المستخدمين الذين يُعتبرون مُخادعين أو مُسيئين؛ هم في صميم استجابات تنظيمية كثيرة تصدرها الحكومات الساعية ليس فقط إلى الحد من التحريض على الكراهية والعنف والتشهير بل أيضاً على نطاق أوسع وبشكل إشكالي، إلى وضع حد للتعبير الذي يُعتبر «كاذباً» أو مضراً بالأمن القومي والدبلوماسية الدولية والنظام الاجتماعي، وغيره (مراجعة الرقم 8 في الجدول 3 أدناه). من ناحية أخرى، تقوم بعض الحكومات بتقديم الدعم للجهات العاملة على ضمان جودة المعلومات أي على مبادرات التحقُّق من الوقائع ومُكافحة التّضليل والدراية الإعلامية والمعلوماتية والمبادرات الصحفية التي تسعى إلى إعلام المستخدمين بطريقة موثوقة وتمكينهم من كشف التّضليل (مراجعة الأرقام 1،2،3،9،10 من الجدول 3 أدناه).

ثانياً، إنّ المبادرات الحكومية التي تركز على شركات التواصل عبر الإنترنت تستهدف السلوكيات الاقتصادية والفنية لهذه الشركات. واستناداً إلى الافتراض الذي مفاده أنّ خوارزميات المنصّات الإلكترونية تسمح بتضخيم المعلومات المُضلّلة لجعلها واسعة الانتشار؛ تُحاول العديد من المبادرات التنظيمية فرض التزامات أكبر على هذه الجهات.

ويُطلب من شركات التواصل عبر الإنترنت، في التدخلات الحكومية الأقل صرامةً، اعتماد التنظيم الذاتي وتقديم المعلومات حول إجراءات وممارسات تنظيم المُحتوى والإعلانات السياسية. أما في الإجراءات التنظيمية الأكثر صرامةً، فيُطلب من منصات وسائط الإنترنت، بشكل رسمي أو غير رسمي، تخفيض مستوى الأولوية الممنوحة لأنواع معينة من المُحتوى والمواقع وحجبها وإلغائها، كما ويُطلب منها إلغاء انتساب بعض المستخدمين (مراجعة الأرقام 5،6،7 في الجدول 3 أدناه).

إلى حدٍّ ما، وإن لم تكن هدفاً مباشراً في كثير من الأحيان، قد تندرج صناعة الإعلان في هذه الفئة، حيث يرى بعض صناع السياسات أنّ نموذج الإعلانات التجارية عبر الإنترنت يسمح بتمويل عمليات التّضليل بشكل غير مباشر (مراجعة الرقم 7 في الجدول 3 أدناه).

أما الطّرف الثالث الذي تُركّز عليه الإجراءات الحكومية التي تتصدى للتضليل فهُم: الصحفيون ووسائل الإعلام. حيث تقوم العديد من الاستجابات التنظيمية إما من حيث تصميمها أو عن غير قصد، بتجريم الصحفيين وناشري الأخبار لنشر وتوزيع المعلومات الخاطئة، على الرغم من الحماية الدولية الضامنة لحُريّة الصحافة. وهذا يدلّ على الحاجة إلى محاذير لحماية الصحفيين (مراجعة الرقم 8 في الجدول 3). في المقابل، وكما ذُكر أعلاه، هناك بعض التدخلات التي حفّزت الاستثمار في الصحافة المستقلة والتعاون بين الوكالات الإخبارية والمجتمعات بهدف تعزيز الدراية الإعلامية والمعلوماتية ومبادرات التحقُّق من الوقائع من قبل أطراف ثالثة (مراجعة الأرقام 1،2،10 في الجدول 3)، كجزء من الإقرار بدور وسائل الإعلام في مواجهة المعلومات المُضلّلة.

وأخيراً، تستهدف بعض الاستجابات الحكومية الفاعلين السياسيين (بما في ذلك الأحزاب السياسية) أنفسهم، من خلال مُطالبتهم بتلبية التزامات جديدة تتعلق بشفافية الحملات السياسية على الإنترنت، من خلال تصنيف الإعلانات السياسية على سبيل المثال (مراجعة الرقمين 3 و7 في الجدول 3 [125])، و/أو عن طريق تعزيز مساعي التحقُّق من الوقائع خلال الفترات الانتخابية (مراجعة الرقمين 1 و7 في الجدول 3).

2.1.5 من هي الجهات التي تُحاول مُساعدتها الاستجابات التشريعية وما قبل التشريعية والسياساتية؟

إنّ الاستجابات للتضليل المتخذة على مستوى الدولة تستهدف جميع الجهات المعنية من مستخدمين نهائيين (أفراد ومجتمعات وجماهير.. إلخ)، ومنصات إلكترونية ومُعلنين وصحافيين ومؤسسات إخبارية وسياسيين وأحزاب سياسية. كما وتستهدف أيضاً الجهات الفاعلة المحلية والأجنبية التي تُعتبر سيئة النيّة. وتسعى هذه التدخّلات التنظيمية إلى ردع ما تراه تعبيراً مُسيئاً في إطار المعلومات المُضلّلة –وهي محور هذه الدراسة–، وذلك عن طريق السياسات والقوانين. وهدفها استخدام تدابير «سلبية» (أي مُقيّدة) لحماية المجتمع وحقه في الحصول على المعلومات عن طريق تقييد وجود المعلومات المُضلّلة المدمرة والضارة. من ناحيةٍ أخرى، تسعى التدابير «الإيجابية» (أي التمكينية) إلى ترسيخ الحق في حُريّة التّعبير عبر تحسين البيئةً المحيطة بالمعلومات من خلال برامج تعزيز الدراية الإعلامية والمعلوماتية وتخصيص الأموال لجهات التحقُّق من الوقائع ولوسائل الإعلام و/أو الخطاب المضاد. ومع ذلك، فإنّ تفسيرات كل دولة لمسألة الحقوق والمسؤوليات لا تتماشَى دائماً مع قصد الأُطُر القانونية والمعيارية الدولية الرامية إلى دعم حُريّة التّعبير.

وتسعى التدابير «السلبية» إلى تقييد أي مُحتوى أو سلوك تعتبره السُّلطات مخادعاً أو مسيئاً بطرق متنوعة. وهي تركّز في المقام الأول على تنظيم الخطاب العام، بحجة الحد من الضرر الذي قد يلحق بالآخرين، ولضمان الصحة العامة ولتأمين لدفاع والأمن ولكن أيضاً، في بعض الأحيان، لتحقيق مكاسب سياسية.

[125] يُرجَى مراجعة القسمين 4.1 و7.1.

ومن المعروف أنّ التدخلات التي تحد من حُريّة حقوق التّعبير قد تنزلق بشكل خَطِر. لذلك، تتطلّب المعايير الدولية أن تكون منصوصةً في القانون، وأن تكون مشروعةً ومتناسبةً، وذات ضرورة مثبتة. كما ويُطلب أن تكون الوسيلة الأقل تقييداً لتحقيق الهدف المعلن. وإذا اعتُمدت في حالات الطوارئ، لابد أن تكون أيضاً محدودةً في الزمن.

من ناحيةٍ أخرى، تهدف التدابير «الإيجابية» التي تُركّز على المستخدمين، على الأقل جزئياً، إلى تعزيز الدراية الإعلامية والمعلوماتية وضمان تمكين المستخدمين من خلال المُحتوى الذي ينفذون إليه عبر الإنترنت. وبالمثل، تمكّن هذه التدابير وسائل الإعلام وتسمح لها بالتحقيق والتحقّق ونشر وتعميم المعلومات الهادفة إلى تحقيق المصلحة العامة. وفي ما يتعلق بعوامل التحفيز، تركز الإجراءات الحكومية في المقام الأول على تشجيع الجهات الفاعلة الأخرى لمواجهة التّضليل، ولكنها تستخدم أيضاً قوة الإكراه القانونية ضد الجهات الناشطة في «صناعة» التّضليل. ونظرية التغيير التي يقوم عليها هذا النوع من الاستجابات تعتمد على الجهات والأمور المستهدفة:

- بالنسبة إلى المستخدمين، يُفترض أنه من الممكن الحد من الخطاب المسيء من خلال تطبيق تدابير عقابية، مثل الغرامات والقبض على الأشخاص. وبالتالي يُفترض إحداث التغيير من خلال زيادة حجم المعلومات الموثوق بها والنفاذ إليها، كما ومن خلال رفع الوعي بين المواطنين وتطبيق برامج الدراية الإعلامية والمعلوماتية المصممة لتلقيح «القطيع» ضد التّضليل، بحيث يصبح المستخدمون أكثر قدرةً على فهم وإدارة عملية إنتاج المُحتوى الخاص بهم وتداوله واستخدامه.
- والنسبة إلى شركات التواصل عبر الإنترنت/وشركات العلاقات العامة والإعلان، تركّز نظرية التغيير على دور القوانين والسياسات في الحد بشكل مباشر أو غير مباشر من الحوافز الاقتصادية والسياسية التي تغذي التّضليل. ويستند هذا أيضاً على الافتراض الذي مفاده أنّ مصلحة الشركات المعنية تكمن في إحباط الفاعلين الذين يستغلّون الفرص التي تتيحها التكنولوجيا ونماذج العمل المعاصرة. وفي بعض الحالات، يكون الهدف هو التحكّم بتدفق المعلومات عن طريق ضمان أن تقوم الشركات بالاستفادة بشكل أفضل من التكنولوجيا على غرار الذكاء الاصطناعي لمعالجة القضايا على نطاق واسع.
- أما بالنسبة إلى الصحفيين وناشري الأخبار، فعلى غرار المستخدمين، تفترض نظريّة التغيير أنه بالإمكان الحد من نشرهم للمعلومات «الخاطئة» وللخطاب المسيء (والذي بطريقة إشكالية قد يكون محط انتقاد قوي كمنتج للصحافة المستقلة) من خلال تدابير عقابية، مثل الغرامات والرقابة والقبض على الأشخاص. وينتج عن ذلك افتراض مترابط يتماشى مع القانون الدولي لحقوق الإنسان، يؤكد على أن التغيير يمكن أن يحدث من خلال دعم الصحافة المستقلة، حيث أنّ توفير معلومات واقعية وقابلة للتدقيق بهدف تحقيق المصلحة العامة هو شرط مسبق لتعزيز الديمقراطية والتنمية المستدامتين.
- وبالنسبة إلى السياسيين، تفترض نظرية التغيير التي ترعى التدخلات التنظيمية ذات الصلة أن الحملات السياسية غير المنظمة على الإنترنت، يمكن أن تخضع لقواعد جديدة أو مُحدَّثة تُناسب البيئة الرقمية. وإنّ التدقيق خلال الفترات الانتخابية، من خلال اعتماد الشفافية في الإعلانات السياسية وتعزيز التحقُّق من الوقائع، يُعتبر حافزاً يثني المرشحين عن استخدام التّضليل كاستراتيجية للتواصل.

وقـد خُصّـص القسـم الفرعـي 5.1.6 أدنـاه لمناقشـة صحـة هـذه التصـورات المتعلقـة بأسـباب التدخـلات وتأثيـر نتائجها.

3.1.5 ما هي مُخرجات الاستجابات التشريعية وما قبل التشريعية والسياستية؟

إنّ مخرجات الاستجابات المتخذة على مستوى الدولة هي تقارير صادرة عن التحقيقات ووثائق سياسية (وبحوث تدعم عملية تطوير السياسات)، ومشاريع قوانين وتشريعات، وأحكام قضائية منشورة. وفي الحالات التي تتخذ الحكومة الإجراءات، فإن النواتج تشمل أيضاً التدبير المتخذ المحدد، مثل الغرامات أو القبض والاحتجاز أو الحملات لمُكافحة ما تعتبره السُّلطات معلومات مضللة، أو إيقاف خدمة الإنترنت. وفي حال التدابير الإيجابية، يتم تخصيص الموارد واتخاذ الخطوات الهادفة إلى بناء القدرات على غِرار برامج الدراية الإعلامية والمعلوماتية، وتطبيق أنظمة لتعزيز النفاذ إلى المعلومات، وتعزيز وتقوية وسائل الإعلام، وما إلى ذلك.

4.1.5 من هي الجهات الفاعلة الرئيسية ومن يموّلها؟

عـادةً، تقـوم الدولـة بتمويـل التدابيـر التشـريعية ومـا قبـل التشـريعية والسياسـتية، ولكـن فـي بعـض الحـالات -مثـل شـركات التواصـل عبـر الإنترنـت - تقـوم كيانـات خاصـة بدفـع تكاليـف التنفيـذ كمـا فـي حـال الامتثـال لمتطلبـات الشـفافية فـي مجـال الإعـلان السياسـي. وهـذا يُطبّـق على العديـد مـن المؤسسـات التجاريـة فـي أكثـر مـن قطـاع؛ التـي عليهـا أن تلتـزم بالتشـريعات والسياسـات الراميـة إلـى حمايـة المصلحـة والسـلامة العامتيـن كجـزء مـن تكاليـف ممارسـة الأعمـال التجاريـة. وفـي الوقـت نفسـه، يمكـن للـدول تمويـل وتنفيـذ حملاتهـا الخاصـة لمُكافحـة التّضليـل، أو برامجهـا الخاصـة لتعزيـز الدرايـة الإعلاميـة والمعلوماتيـة.

ويغطي القسـمان 5.1 و5.2 من هذا الفصل العديد من الاستجابات الحكومية في جميع أنحاء العالم، اي 117 استجابةً عبـر 61 دولـةً ومنظمـةً حكوميـةً دوليـةً. وفـي حيـن كان الهـدف إثبـات وجود مجموعـة مـن الخبرات؛ لا شـك فـي أن الهفوات لا مفرَّ منها. ومعظـم هـذه المبـادرات السياسـية هـي حديثـة للغاية، والعديد منها خضع لتغيير ومراجعة منذ كتابة التقرير. وبالإضافة إلـى ذلك، قد تتحول التحقيقات إلى مقترحات تشـريعية، في حيـن أن بعـض المقترحـات التشـريعية قد لا تُعتمد، وقد تظهـر لوائـح تنظيمية جديدة أو قد تعتمد التعديلات عليهـا.. إلـخ. ولذلك ينبغـي أن ينظـر إلـى هـذا المسـح كأداة متغيـرة. ويبين الجـدول أدنـاه أيضاً فئات عامة، ولا يقدم تفصيلاً لمسـائل كتجريم التّضليل ضمن فئة الاسـتجابات التشـريعية.

ويضم هذا القسم ملخصاً لنتائج البحث. ولتحليل كل حالة؛ لذا يُرجى مراجعة الملحق أ[126].

إنّ الأرقـام الـواردة فـي الجـزء العلـوي مـن الجـدول أدنـاه تعكـس الاسـتجابات للمعلومـات المُضلّـلة كمـا تـرد فـي التصنيـف العـام المعتمـد فـي هـذه الدراسـة، وتبيـن الروابـط بيـن الاسـتجابات التشـريعية/السياسـتية والاسـتجابات الأخـرى.

1. استجابات الرصد/التحقُّق من الوقائع.
2. الاستجابات الاستقصائية.

[126] يُرجَى مراجعة الملحق أ.

3. الحملات الوطنية والدولية لمُكافحة التّضليل.
4. الاستجابات الخاصة بالانتخابات.
5. الاستجابات التنظيمية.
6. الاستجابات التقنية/الخوارزمية.
7. الاستجابات الاقتصادية.
8. الاستجابات الأخلاقية والمعيارية.
9. الاستجابات التثقيفية.
10. استجابات التمكين وتصنيف المصداقية.

نوع السياسة		الاستجابة السياسية	1. استجابات الرصد/التحقُّق من الوقائع	2. الاستجابات الاستقصائية	3. الحملات الوطنية والدولية لمُكافحة التّضليل الحملات	4. الاستجابات الخاصة بالانتخابات	5. الاستجابات التنظيمية	6. الاستجابات التقنية/الخوارزمية	7. الاستجابات الاقتصادية	8. الاستجابات الأخلاقية والمعيارية	9. الاستجابات التثقيفية	10. استجابات التمكين وتصنيف المصداقية
التحقيقات وفرق العمل والمبادئ التوجيهية		الجهات الفاعلة: رابطة أُمم جنوب شرق آسيا (أو رابطة أمم جنوب شرق آسيا)، أستراليا، بلجيكا، البرازيل، كندا، مجلس أوروبا، الدنمارك، إستونيا، الاتحاد الأوروبي، الهند،إندونيسيا، اللجنة الدُّوليّة الكبرى، آيرلندا، إيطاليا، اليابان، المكسيك، هولندا، نيوزيلندا، منظمة الدول الأمريكية، جنوب أفريقيا، جمهورية كوريا، إسبانيا، السويد، أوكرانيا، المملكة المتحدة، الولايات المتحدة										
	1	الإعلان المشترك لوزراء رابطة أمم جنوب شرق آسيا المسؤولين عن الإعلام	x		x					x	x	
	2	فريق عمل أستراليا لضمان الانتخابات				x	x	x	x		x	
	3	اللجنة الدائمة المشتركة للبرلمان الأُسترالي حول المسائل الانتخابية: الديمقراطية والتّضليل	x			x				x	x	
	4	فريق الخبراء والمنصة التشاركية في بلجيكا	x								x	x
	5	المحكمة الانتخابية العليا في البرازيل	x		x	x	x					
	6	تقرير اللجنة البرلمانية في كندا حول «الديمقراطية المهددة»				x	x	x	x		x	
	7	مبادرة المواطن الرقمي في كندا			x				x	x	x	x
	8	البروتوكول الكندي حول الحوادث الانتخابية الخطيرة			x	x				x	x	
	9	دراسة «اضطراب المعلومات» لمجلس أوروبا	x	x			x	x	x		x	x
	10	خطة العمل الخاصة بالانتخابات في الدنمارك	x	x		x						
	11	رابطة الدفاع السيبراني في إستونيا						x				
	12	مدونة الممارسات وخطة العمل بشأن التّضليل في الاتحاد الأوروبي	x		x	x	x	x	x		x	x
	13	مدونة الأخلاقيات حول منصات التواصل الاجتماعي في الهند				x	x	x			x	

نوع السياسة		الاستجابة السياسية	1. استجابات الرصد/التحقُّق من الوقائع	2. الاستجابات الاستقصائية	3. الحملات الوطنية والدولية لمُكافحة التّضليل الحملات	4. الاستجابات الخاصة بالانتخابات	5. الاستجابات التنظيمية	6. الاستجابات التقنية/الخوارزمية	7. الاستجابات الاقتصادية	8. الاستجابات الأخلاقية والمعيارية	9. الاستجابات التثقيفية	10. استجابات التمكين وتصنيف المصداقية
التحقيقات وفرق العمل والمبادئ التوجيهية	14	غرفة العمليات والحملات «وقف الخدع» Stop Hoax في إندونيسيا	x	x	x	x	x			x	x	
	15	اللجنة الدولية الكبرى حول «التّضليل و«الأخبار الزائفة»				x	x	x	x			
	16	المجموعة المشتركة بين الإدارات في آيرلندا حول «أمن العملية الانتخابية والتّضليل»				x	x				x	
	17	حملة «كفى خدعاً» Enough-with-the-Hoaxes وبوابة «الزر الأحمر» Red Button	x	x		x					x	
	18	مجموعة الدراسة حول الخدمات على المنصات في اليابان				x	x					
	19	المعهد الانتخابي الوطني في المكسيك	x			x	x				x	x
	20	حملة واستراتيجية «تَحَلَّ بحس نقدي» في هولندا	x	x		x					x	x
	21	التحقيق البرلماني في انتخابات نيوزيلندا لعامي 2016 و2017				x	x	x	x		x	
	22	دليل منظمة الدول الأميركية حول حُريّة التّعبير والتّضليل خلال الانتخابات	x	x		x	x	x	x		x	x
	23	أرشيف إعلانات الأحزاب السياسية والشكاوى المتعلقة بالمعلومات المُضلّلة الرقمية في جنوب أفريقيا				x					x	x
	24	فريق العمل الحزبي في جمهورية كوريا	x	x								
	25	وحدة التهديدات المختلطة في حكومة إسبانيا	x	x		x						
	26	التحقيق في تطوير هيئة الدفاع النفسي في السُّويد									x	x
	27	مبادرة «تعلم التمييز» Learn to Discern في أوكرانيا									x	
	28	تحقيق مجلس العموم في المملكة المتحدة (اللجنة المعنية بالشؤون الرقمية والثقافة والإعلام والرياضة) حول «التّضليل و»الأخبار الزائفة»»	x			x	x	x	x		x	
	29	تحقيق لجنة الشؤون الخارجية التابعة لمجلس العموم في المملكة المتحدة حول حريّة الإعلام العالمي (وموضوعه الفرعي هو المعلومات المُضلّلة)	x	x	x					x		
	30	لجنة مجلس الشيوخ الأميركي حول تحقيق وكالات المخابرات في «حملات التدابير الناشطة والتدخل الروسي في الانتخابات الأميركية لعام 2016»				x						

نوع السياسة		الاستجابة السياسية	1. استجابات الرصد/التحقُّق من الوقائع	2. الاستجابات الاستقصائية	3. الحملات الوطنية والدولية لمُكافحة التّضليل الحملات	4. الاستجابات الخاصة بالانتخابات	5. الاستجابات التنظيمية	6. الاستجابات التقنية/الخوارزمية	7. الاستجابات الاقتصادية	8. الاستجابات الأخلاقية والمعيارية	9. الاستجابات التثقيفية	10. استجابات التمكين وتصنيف المصداقية
المقترحات التشريعية		الجهات الفاعلة: الأرجنتين، تشيلي، فرنسا، ألمانيا، الهند، آيرلندا، إسرائيل، نيجيريا، الفلبين، جمهورية كوريا، سريلانكا، المملكة المتحدة، الولايات المتحدة										
	31	اقتراح قانون لإنشاء لجنة للتحقق من الأخبار الزائفة في الأرجنتين	x			x	x	x			x	
	32	اقتراح تشيلي «لإنهاء ولاية السياسيين المنتخبين بسبب التّضليل»				x						
	33	اقتراح فرنسا حول خطاب الكراهية على الإنترنت					x	x				
	34	تحديث القانون حول شبكات التواصل في ألمانيا					x			x		
	35	التعديلات المقترحة على المبادئ التوجيهية لوسائط التكنولوجيا المعلومات في الهند				x	x					
	36	اقتراح آيرلندا لتنظيم شفافية الإعلانات السياسية على شبكة الإنترنت				x	x					
	37	التعديلات على القانون الانتخابي المقترح في إسرائيل والقوانين حول فيس بوك				x	x					
	38	مشروع قانون للحماية من الزيف والتلاعب على الإنترنت في نيجيريا						x		x		
	39	مشروع قانون الفلبين لمُكافحة المحتويات الكاذبة										
	40	اقتراحات القوانين في جمهورية كوريا				x	x			x		
	41	التعديلات المقترحة على قانون العقوبات في سريلانكا								x		
	42	الورقة البيضاء حول الأذى على الإنترنت في المملكة المتحدة				x	x					
	43	مشروع قانون ولاية تينيسي الأميركية لتسجيل محطة CNN وصحيفة واشنطن بوست كوكلاء «للأخبار الزائفة» تابعين للحزب الديمقراطي			x							x

نوع السياسة		الاستجابة السياسية	1. استجابات الرصد/التحقّق من الوقائع	2. الاستجابات الاستقصائية	3. الحملات الوطنية والدولية لمُكافحة التضليل الحملات	4. الاستجابات الخاصة بالانتخابات	5. الاستجابات التنظيمية	6. الاستجابات التقنية/الخوارزمية	7. الاستجابات الاقتصادية	8. الاستجابات الأخلاقية والمعيارية	9. الاستجابات التثقيفية	10. استجابات التمكين وتصنيف المصداقية
التشريعات المعتمدة		الجهات الفاعلة: الأرجنتين، بنغلاديش، روسيا البيضاء، بنين، البرازيل، بوركينا فاسو، كمبوديا، الكاميرون، كندا، الصين، ساحل العاج، مصر، إثيوبيا، فرنسا، ألمانيا، إندونيسيا، كازاخستان، كينيا، ماليزيا، ميانمار، نيوزيلندا، عُمان، باكستان، الفلبين، الاتحاد الروسي، سنغافورة، تايلاند، فيتنام										
	44	قانون تمويل الأحزاب السياسية في الأرجنتين				x			x		x	x
	45	قانون الأمن الرقمي في بنغلاديش	x				x	x		x		
	46	قانون الإعلام في بيلاروس						x		x		
	47	القانون الرقمي في بنين								x		
	48	القانون الجنائي حول التّضليل الانتخابي في البرازيل				x						
التشريعات المعتمدة	49	القانون الجنائي في بوركينا فاسو					x	x		x		
	50	توجيهات كمبوديا حول مُكافحة الأخبار الزائفة					x	x		x		
	51	القانون الجنائي وقانون الأمن السيبراني والجرائم السيبرانية								x		
	52	قانون تحديث الانتخابات في كندا				x	x					
	53	قوانين مُكافحة الشائعات في الصين					x	x		x		
	54	القانون الجنائي وقانون الصحافة في ساحل العاج								x		
	55	قوانين مصر لمُكافحة الأخبار الزائفة					x	x		x		
	56	قانون اثيوبيا حول المعلومات الخاطئة								x		
	57	القانون الفرنسي ضد التلاعب بالمعلومات				x	x	x				
	58	القانون الألماني لتحسين إنفاذ القانون على الشبكات الاجتماعية					x	x				
	59	قانون المعلومات والمعاملات الإلكترونية في إندونيسيا					x	x		x		
	60	القانون الجنائي في كازاخستان								x		
	61	قانون كينيا حول سوء استعمال الحاسوب والجرائم الالكترونية								x		
	62	قانون ماليزيا لمُكافحة الأخبار الزائفة (مُلغَى)								x		
	63	قانون الاتصالات والقانون الجنائي في ميانمار								x		
	64	تعديل قانون الانتخابات في نيوزيلندا				x			x		x	
	65	القانون الجزائي في عُمان					x	x		x		
	66	قانون باكستان حول مُكافحة الجرائم الإلكترونية					x	x		x		
	67	القانون الجنائي في الفلبين					x	x		x		
	68	تعديلات تتعلق بالأخبار الزائفة على قانون الإعلام الروسي وقانون المخالفات الإدارية					x	x		x		
	69	قانون سنغافورة للحماية من الأكاذيب والتلاعب على الإنترنت					x	x	x			
	70	قانون جرائم الحاسوب في تايلاند					x	x		x		
	71	قانون الأمن السيبراني في فيتنام					x	x		x		

نوع السياسة		الاستجابة السياسية	1. استجابات الرصد/التحقُّق من الوقائع	2. الاستجابات الاستقصائية	3. الحملات الوطنية والدولية لمُكافحة التّضليل الحملات	4. الاستجابات الخاصة بالانتخابات	5. الاستجابات التنظيمية	6. الاستجابات التقنية/الخوارزمية	7. الاستجابات الاقتصادية	8. الاستجابات الأخلاقية والمعيارية	9. الاستجابات التثقيفية	10. استجابات التمكين وتصنيف المصداقية
إنفاذ القانون وتدخلات حكومية أخرى		الجهات الفاعلة: البحرين، بنغلاديش، بنين، وكمبوديا، الكاميرون، الصين، ساحل العاج، مصر، ألمانيا، الهند، إندونيسيا، كازاخستان، لاتفيا، ماليزيا، ميانمار، الاتحاد الروسي، سنغافورة، سريلانكا، تايلاند، أوكرانيا										
	72	البحرين								x		
	73	بنغلاديش				x		x				
	74	بنين								x		
	75	كمبوديا								x		
	76	الكاميرون								x		
	77	جمهورية الصين الشعبية					x	x		x		
	78	ساحل العاج								x		
	79	مصر						x		x		
	80	ألمانيا					x					
	81	الهند						x				
	82	إندونيسيا					x	x		x		
	83	كازاخستان					x	x		x		
	84	لاتفيا						x				
	85	ماليزيا								x		
	86	ميانمار								x		
	87	الاتحاد الروسي						x				
	88	سنغافورة					x					
	89	سيريلانكا						x				
	90	تايلاند								x		
	91	أوكرانيا						x				

الجدول 3. الاستجابات التشريعية وما قبل التشريعية والسياستية (تم مسحها وفقاً للتصنيف المعتمد في هذه الدراسة)

التحقيقات وفرق العمل والمبادئ التوجيهية

مع انتشار المعلومات المغلوطة والمعلومات المُضلّلة أصبحت هذه الظاهرة مصدر قلق متزايد، مما دفع بالعديد من الدول إلى إنشاء فرق عمل متخصصة وإطلاق تحقيقات بهدف رصد حملات التّضليل والتحقيق فيها. وغالباً ما أنشِئت فرق العمل هذه بعد حملات للتضليل اعتُبِرت تهديداً مختلطاً لسلامة الديمقراطية في البلاد أو للأمن السيبراني. والتثقيف هو من الأهداف الأخرى لهذه المبادرات الحكومية التي تتضمّن بمعظمها برامج لتعزيز الدراية الإعلامية والمعلوماتيّة (يُرجى مُراجعة رقم 9 في الجدول 3)، على غِرار استراتيجية هولندا «تحلَّ بالحس الناقد» أو 'Stay Critical' (البند 20 في الملحق أ) أو مبادرة «المواطن الرقمي» Digital Citizen في كندا (البند 7 في الملحق أ). وبالإضافة إلى ذلك، تتضمن 17 مبادرة في هذه الفئة تدابير للتحقق من الوقائع (يُرجَى مراجعة الرقم 1 في الجدول 3). ولابد من الإشارة إلى أنه من بين الدول الثلاثين التي أقامت هذه التحقيقات أو فرق العمل، 21 دولة ضمّنتها عنصراً خاصاً بالانتخابات (يُرجَى مراجعة رقم 4 في الجدول 3)، بما في ذلك لجنة مجلس الشيوخ الأمريكي حول تحقيق المخابرات في التدخل في الانتخابات الأمريكية (البند 30 في الملحق أ)، وفريق عمل أستراليا لضمان الانتخابات (البند 2 في الملحق أ)، والمعهد الانتخابي الوطني في المكسيك (البند 19 في الملحق أ). وتهدف التحقيقات الخاصة بالانتخابات إلى التحقيق في العمليات التشريعية أو منع التدخل فيها. ولأنّ التّضليل على الإنترنت هو ظاهرة جديدة نسبياً، فإنّ معظم المبادرات التي تم تحديدها؛ حديثة ولا تزال قيد التطوير، بما في ذلك المبادرات التنظيمية.

المقترحات التشريعية

إنّ معظم المقترحات التشريعية الحديثة (8 من أصل 13 تم تحليلها) تهدف إلى معالجة التّضليل من خلال التنظيم ومن منظور التزامات المنصّات في ما يخص المعلومات المغلوطة/المُضلّلة أو خطاب الكراهية (مراجعة الرقم 5 في الجدول 3). وهذا هو نطاق اقتراح القانون الفرنسي لمُكافحة خطاب الكراهية على الإنترنت (البند 37 في الملحق أ)، واقتراح آيرلندا لتنظيم شفافية الإعلانات السياسية على الإنترنت (البند 39 في الملحق أ) والتعديلات المقترحة على القانون الانتخابي في إسرائيل و«القوانين المتعلقة بفيس بوك» (البند 40 في الملحق أ). وعلى غرار التحقيقات وفرق العمل، تركز المقترحات التشريعية في بعض الأحيان على الأمور الانتخابية (مراجعة الرقم 4 في الجدول 3)، مثل: اقتراح شيلي لإنهاء ولاية السياسيين المنتخبين بسبب التّضليل (البند 35 في الملحق أ). كذلك تعمد بعض المقترحات التشريعية الأخرى إلى تجريم فعل نشر المعلومات المُضلّلة (مراجعة الرقم 8 في الجدول 3). ولكن قد يؤدي ذلك إلى خطر يحذر منه الناشطون في مجال حقوق الإنسان في عدة مناسبات؛ وهو أن يُستخدم هذا التجريم ضد الصحفيين المستقلين المنتقدين للسلطات.

التشريعات المعتمدة

وفقا لهذا التقرير البحثي، في آذار/مارس 2020، كان 28 بلداً على الأقل قد اعتمد تشريعات متعلقة بالتّضليل، إما عن طريق تحديث تشريعات قائمة أو عن طريق إصدار تشريعات جديدة. ويمتد نطاق التشريع من وسائل الإعلام والقوانين الانتخابية إلى الأمن السيبراني والقوانين الجنائية. وتستهدف اللوائح التنظيمية مُرتكبي ما تعتبره السُّلطات معلومات مضللة (خاصةً الأفراد والمؤسسات الإعلامية) أو تحمّل شركات التواصل عبر الإنترنت مسؤولية تنظيم أو إزالة مُحتوى معيّن، مثل: القانون الألماني حول شبكات التواصل (البند 61 في الملحق أ).

في بعض الحالات، وخاصةً حيث يتم تعريف التّضليل بشكل واسع أو حيث يتم إدراج الأحكام المتعلقة بالتّضليل في القوانين الجنائية، يكون خطر الرقابة كبيراً.

إنفاذ القانون وتدخُّلات الدول الأخرى

لقد استخدمت الدول قوانين قائمةً أو أُقرت مؤخراً لتبرير عدد من التدخلات التي اعتمدتها للحد من التّضليل. وتشمل هذه الإجراءات غرامات واعتقالات أو إغلاق مواقع على الإنترنت. ويستهدف الإنفاذ الأفراد وأحياناً الصحفيين والناشطين ووسائل الإعلام الحكومية الأجنبية التي تُعتبر ناشرةً للمعلومات المُضلّلة (مثال على ذلك: قيام «لاتفيا» التي قامت بإغلاق موقع على شبكة الإنترنت مرتبط بحكومة أخرى (البند 88 في الملحق أ)؛ أو شركات التواصل عبر الإنترنت التي تُعتبر مسؤولةً عن الانتشار الواسع للمعلومات المُضلّلة (يُرجَى مراجعة الغرامات التي فرضتها ألمانيا على فيس بوك (البند 84 في الملحق أ). وقد أشار عدد من منظمات حقوق الإنسان إلى عدد من الاعتقالات التعسفية التي استغلت مُكافحة التّضليل للحد من حُريّة التّعبير. كما وقد عمدت بعض الحكومات إلى إغلاق شبكة الإنترنت بحجة منع انتشار المعلومات المُضلّلة على الرغم من أنّ هذه التدابير تضع قيوداً واضحةً (تكون شاملةً أكثر أو اقل من اللزوم) على النفاذ إلى معلومات، ما كانت لتكون متاحةً لأفراد المجتمع لولا هذه التدابير.

5.1.5 كيف يتم تقييم هذه الاستجابات؟

يصعب تقييم أثر المعلومات المُضلّلة بشكل شامل، كما ويصعب تقييم فعالية القوانين المعتمدة في التصدي لها. ومع ذلك، تعتمد بعض الشركات تدابير ضد التّضليل على الإنترنت (التحذير والمراجعة والانتقائية والحظر) يمكن اعتبارها معايير لتقييم تطبيق القانون. على سبيل المثال، كوسيلة لتقييم تطبيق القانون الألماني حول شبكات التواصل، تُفيد المنصات كل ستة أشهر عن الإجراءات التي اتخذتها إزاء المُحتوى الذي يبلّغ عنه المستخدمون. وكان ذلك، جزئياً، السبب الذي دفع بالحكومة الألمانية إلى اقتراح تحديثات على القانون بعدما تبيّن أن هناك نقص في التبليغ عن عدد من الشكاوى التي وردت (Pollock, 2019). وكان من المتوقّع أن يجهز في منتصف عام 2020. وكان من المتوقع أن يوضع نص ثان يهدف إلى مراجعة القانون مع التركيز على إدارة الشكاوى على المنصات (German BMJV, 2020a; German BMJV, 2020b) (البنود 37 و61 و84 في الملحق أ).

وقد اتضح أيضاً أنه يصعب تطبيق بعض القوانين على أرض الواقع. على سبيل المثال، بعد اعتماد القانون الفرنسي لمُكافحة التلاعب بالمعلومات (البند 60 في الملحق أ)، قامت الجهات المعنية والمرشحون السياسيون بإلقاء الضوء على محدودية هذا القانون. بالإضافة إلى ذلك، قام تويتر بحظر حملة إعلامية حكومية رسمية تهدف إلى تشجيع الناخبين على التصويت، بحجة الامتثال للقانون (LeFigaro, 2019). وبالنسبة إلى العديد من البلدان الصغيرة في العالم، يصعب من الناحية العملية تطبيق القوانين على شركات دولية ليست لديها أعمال كثيرة أو وجود فعلي في نطاق الولاية القضائية الوطنية.

ويمكن للحكومات والبرلمانات والمحاكم تقييم، وإذا لزم الأمر، إعادة النظر في التشريعات والسياسات القائمة وتعديلها. على سبيل المثال، تم الطعن في دستورية قانون كينيا حول سوء استعمال الحاسوب والجرائم الإلكترونية الصادر عام 2018 في المحكمة وكان من المتوقع صدور حكم في أوائل عام 2020 (البند 64 في الملحق أ). وفي عام 2018 أقرت ماليزيا قانون مُكافحة الأخبار الزائفة. ولكن بعد أن تم

تغيير الحكومة، أُلغي القانون على أساس أن القوانين القائمة (قانون العقوبات، قانون إثارة الفتنة، قانون المطابع والمطبوعات والاتصالات والوسائط المتعددة) تعالج مسألة التّضليل (البند 65 في الملحق أ).

ويمكن للجهات الفاعلة من غير الدول ممارسة الضغط لتغيير السياسات من خلال نشر تقييماتها ومواقفها الخاصة بشأن المبادرات التنظيمية. والحقيقة أنّ العديد من منظمات المجتمع المدني يقوم في الواقع بتقديم بعض التقييمات. وهذا ما تفعله أيضاً منظمات الأمم المتحدة مثل: المقرر الخاص للأمم المتحدة حول حُريّة الرأي والتّعبير.[127]

6.1.5 دراسة حالة بشأن الاستجابة: التّضليل المتعلق بجائحة فيروس كورونا

لقد دفعت جائحة فيروس كورونا بالدول إلى اتخاذ سلسلة من الإجراءات لمنع ومعاقبة أعمال التّضليل التي قد تشكل خطراً على الحياة (Posetti & Bontcheva, 2020a).[128] ففي جميع أنحاء العالم قامت البرلمانات والحكومات والهيئات الناظمة بتعديل أو إصدار قوانين أو لوائح تنظيمية تسمح بمحاكمة من يختلق وينشر معلومات مضللة، مع أحكام بالسجن تصل إلى خمس سنوات (Quinn, 2020). وفعلياً قامت هذه القوانين بتجريم اختلاق أو مشاركة المعلومات التي تعتبر خاطئةً أو مضللةً و/أو تتعارض مع الرسائل الرسمية التي تنشرها الحكومات حول فيروس كورونا. ومن بين هذه التدابير مراسيم لها طابع الطوارئ أعطت للقادة السياسيين سلطات واسعة جديدة. كذلك طُبقت قوانين الطوارئ القائمة على المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا للسماح باعتقال مرتكبي جرائم مرتبطة بالتّضليل وتغريمهم وسجنهم، كما في جنوب أفريقيا (South African Government, 2020). وعلى سبيل المثال، في كانون الثاني/يناير 2020، استوقفت لجنة الاتصالات والوسائط المتعددة في ماليزيا (2020) أربعة أشخاص اشتُبه في قيامهم بنشر أخبار كاذبة حول جائحة فيروس كورونا وذلك بموجب المادة 233 من قانون الاتصالات والوسائط المتعددة.

وتهدد هذه التدابير باصطياد الصحافة المشروعة في شباكها. (UK Delegation to the OSCE, 2020) ففي بعض البلدان، أُلقي القبض على كُتّاب الصحافة المستقلة وتم القبض عليهم، أو ترحيلهم بموجب هذه القوانين في سياق استجابة الدول لما تعتبره معلومات كاذبة (Simon, 2020; Eljechtimi, 2020). كذلك، تأثرت حقوق حُريّة التّعبير على نطاق أوسع بفعل تطبيق تدابير الطوارئ بطرق تسمح بمعالجة المخاطر الصحية والسلامة العامة بشكل عاجل، فضلاً عن حالات تقييد الوصول إلى المعلومات الرسمية. وفي كثير من الأحيان لم يكن للقيود من تبرير ولم تكن تتماشى مع المعايير التي تفرض عليها أن تكون قانونيةً وضروريةً ومحدودةً في الزمن ومتناسبةً مع الغرض المنشود.

وتتضمن الأنواع الأخرى من الاستجابات السياسية دعم وسائل الإعلام الإخبارية باعتبارها حصناً منيعاً ضد التّضليل. وفي ضوء الأثر السلبي للأزمة على القطاع الإعلامي (Tracy, 2020)، والإقرار بالقيمة الاجتماعية للحفاظ على الوكالات الإخبارية، اتخذت بعض الدول هذا النوع من الإجراءات.

....

127 https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/LegislationAndPolicy.aspx

128 يُرجَى مراجعة قواعد البيانات حول انتهاكات حُريّة التّعبير المتعلقة بالاستجابات الهادفة إلى التصدي للمعلومات المُضلّلة حول جائحة فيروس كورونا (على سبيل المثال: الأخبار الزائفة) التي يُديرها المعهد الدولي للصحافة: https://ipi.media/covid19-media-freedom-monitoring/and Index on Censorship https://www.indexoncensorship.org/disease-control/

على سبيل المثال:

- أقرت كندا كإجراء عاجل تخفيض الضرائب المفروضة على وسائل الإعلام، وتمويل إعلانات توضع خصيصاً على مواقع الوكالات الإخبارية (Canadian Heritage, 2020).
- كذلك أقرت الدنمارك وبلجيكا والمجر وإيطاليا حُزم مساعدات حكومية أو إعفاءات ضريبية لدعم وسائل الإعلام الإخبارية وأصحاب العمل في القطاع الإعلامي (UNI Global Union, 2020).
- ازدادت الدعوات (Aaron, 2020) لاعتماد هذا النوع من الاستجابات السياسية، والإصرار على ضمان الشفافية والنزاهة والاستقلال في آليات الدعم هذه. كذلك كانت هناك دعوات لمساعدة وسائل إعلام الخدمة العامة (Public Media Alliance, 2020).
- قامت بعض المنظمات غير الحكومية بتخصيص أموال للتغطية الإعلامية لجائحة فيروس كورونا مع دعم من الدول (UK Government, 2020)

7.1.5 التحديات والفُرص

إنّ التحدي الأساسي هو وتيرة التغيُّر التكنولوجي الذي يفوق أي إجراء تنظيمي سرعة. ويمكن اعتماد لغة عامة للتغلّب على هذا التحدي، ولكن هذا يفسح المجال أمام تفسيرات تسمح بالتنفيذ الانتقائي أو توسيع نطاق هذه الإجراءات، كما وتفسح المجال أمام جهات فاعلة أخرى لإيجاد ثغرات وتجنب الامتثال.

وثمة تحدّ آخر هو أنه على الرغم من أنّ لمُعالجة التّضليل على المستوى الوطني مزايا واضحة، حيث تُصمَّم المبادرات الحكومية خصيصاً لتُناسب سياقاً سياسياً واجتماعياً محدداً، لا يمكن تطبيق ذلك على المستويات ما فوق الوطنية. فهذه هي حال التدابير التي تستهدف شركات التواصل عبر الإنترنت والتي تعمل على الصعيد العالمي بشكل خاص. في الوقت نفسه، قد يكون من الصعب على الجهات الفاعلة العالمية تنفيذ اللوائح التنظيمية الوطنية المتباينة في سياق تدفُّقات المعلومات الدولية المتشابكة.

وتقضي بعض التدابير المبيّنة في هذا الفصل بتحديث التشريعات القائمة للتقليص من سوء استخدام حُريّة التّعبير، ولتنظيم الانتخابات، من أجل الحد من تأثير التّضليل على قدرة الناخبين على اتخاذ قرارات مُستنيرة ومُتعقّلة. وحيث يتضمن التشريع القائم حماية حُريّة التّعبير والمشاركة الديمقراطية، قد يحول تحديث أو تكييف هذه القوانين لضمان القدرة على تطبيقها على التّضليل على الإنترنت دون الاستعجال في إقرار تشريعات قد تكون لا تحترم المعايير الدولية لحقوق الإنسان.

وغالباً ما تُعتمد التشريعات من دون نقاش أو شفافية كافيين تحت ضغط الرأي العام وتحت ضغط الوقت، خاصةً في الفترة التي تسبق الانتخابات وفي سياق الأزمات والجوائح الصحية الكُبرى مثل جائحة فيروس كورونا. ومن الجدير بالذكر أنه تم الطعن في بعض التشريعات المقترحة والمعتمدة أمام المحاكم، في حين تم تعديل مُسوّدات القوانين وغيرها من القوانين أو سحبها رداً على هذه التطورات.

وعلاوةً على ذلك، في حين تُحاول بعض الحكومات تحديث البيئة التنظيمية لمواجهة التّضليل في العصر الرقمي بحسن نية، يحاول البعض الآخر السيطرة على تعبير المواطنين من خلال اعتماد تصنيفات جديدة لما يُعتبر خطاباً غير قانوني، أو توسيع القوانين القائمة لمعاقبة الخطاب الشرعي. والمفارقة هنا هي أنّ الحكومات التي تسعى لتتحكم بالتّعبير لتحقيق مكاسب سياسية؛ تُحاول إضفاء الشرعية على إجراءاتها من خلال إسنادها إلى لوائح تنظيمية وقوانين مُكافحة التّضليل وخطاب الكراهية. بعبارةٍ أخرى، قد تُستخدم الاستجابات للمعلومات المُضلّلة لفرض رقابة على التّعبير الشرعي (أو تُبرر لغرض) مما يُفسح المجال أمام انتشار المعلومات المُضلّلة الرسمية من دون رادع.

وقد أثار هذا بشكل متزايد مخاوف منظمات حقوق الإنسان في جميع أنحاء العالم، التي أشارت إلى أنّ مثل هذه القوانين قد أدت إلى قبض تعسُّفي بحق صحافيين ونشطاء (Human Rights Watch, 2018b; Posetti & Bontcheva, 2020a; Posetti & Bontcheva, 2020b). ولا شكّ في أنه من الصعب وضع لوائح وقوانين مناهضة للتضليل والحفاظ على حقوق حُريّة التّعبير المكرّسة دولياً في الوقت نفسه إلا أنه لابد من إلقاء الضوء على الفرص المتاحة. على سبيل المثال، عندما يتم تمكين الصحافة المستقلة النقدية لكي تتصدى للتضليل؛ يسمح ذلك بمساءلة الحكومات والشركات الخاصة بفعالية أكبر، كما ويسمح ذلك بتقييم العمل السياسي وتغييره حسب الاقتضاء.

ولابدّ من الإشارة إلى أنّ العديد من الاستجابات التشريعية والسياسية تضع مسؤولية اتخاذ الإجراءات على شركات التواصل عبر الإنترنت (خاصةً اللاعبين العالميين الكبار)، وتحملها مسؤولية الانتشار الواسع للمعلومات المُضلّلة على الإنترنت. ولكن يتم ذلك أحياناً من دون نقاش وشفافية كافيين حول طريقة تطبيق الشركات للتدابير وطريقة الحد من المخاطر التي لا مفرَّ منها. ويتعيّن على الشركات الخاصة بشكل متزايد تنفيذ السياسات الحكومية المتعلقة بالتّضليل، وبالتالي تحديد معالم الخطاب المقبول وغير المقبول في سياق تطبيقها لهذه السياسات. وغالباً ما يتم ذلك من دون أن يكون هناك إمكانيات كافية للتعويض على المستخدمين.

والفرص متاحة لإقامة توازن بين المقاربات التقييدية والإجراءات التمكينية. فبدلاً من اختراع فئات جديدة من الجرائم القائمة على أساس التّعبير، أو تقييد الوصول إلى الإنترنت، هناك استجابات تشريعية وسياسية تساعد على ضمان انتشار المعلومات الصحيحة على الإنترنت بدلاً من المعلومات المُضلّلة. وفي جميع الحالات، من الأفضل تعميم عملية تقييم الأثر على حقوق الإنسان وعلى أشكال التّعبير المشروعة على وجه الخصوص. وهذا يعني اقتراح واعتماد وتنفيذ استجابات على مستوى الدولة للتصدي لمظاهر التّضليل الرقمي.

8.1.5 توصيات تتعلق بالاستجابات التشريعية وما قبل التشريعية والسياساتية

استناداً إلى تقييم الاستجابات التشريعية وما قبل التشريعية والسياساتية للتضليل بطريقة تستند إلى البحوث كما ذُكر أعلاه (وفي الملحق) نتقدم بالتوصيات التالية ليتم النظر فيها فردياً من قبل كل دولة:

- استعراض وتكييف الاستجابات للتضليل بهدف مواءمتها مع المعايير الدولية لحقوق الإنسان (وخاصةً حُريّة التّعبير، بما في ذلك النفاذ السّلس إلى المعلومات، وحقوق الخصوصية)، والسماح بالمتابعة والرصد والتقييم.
- تطوير آليات رقابة مستقلة وتقييم لفعالية التشريعات والسياسات واللوائح ذات الصلة.
- تطوير آليات مستقلة للرقابة ولتقييم ممارسات شركات التواصل عبر الإنترنت ووفائها بالتزاماتها القانونية في التصدي للتضليل.
- تجنّب تجريم التّضليل لكي لا يتم اصطياد الصحافة المشروعة وإعلام المصلحة العامة في شباك قوانين «الأخبار الزائفة».
- تجنُّب اعتماد إغلاق الإنترنت وفرض القيود على وسائل التواصل الاجتماعي كآليات لمُكافحة التّضليل.
- ضمان أن يكون أي تشريع يهدف للتصدي لأزمات التّضليل، على غِرار أزمة المعلومات المُضلّلة في ما يخص فيروس كورونا، ضروري ومتناسب ومحدود في الزمن.
- دعم الاستثمار في تمكين وسائل الإعلام المستقلة، بما في ذلك وسائل الإعلام المجتمعية ووسائل إعلام الخدمة العامة، لمواجهة الآثار الاقتصادية لجائحة فيروس كورونا التي تهدد استدامة الصحافة في جميع أنحاء العالم.

2.5 الحملات الوطنية والدولية لمُكافحة المعلومات المُضلّلة

المؤلفون: تريشا ماير، كلارا هانو، جولي بوزيتي

يُسلّط هذا القسم الضوء على أمثلة مأخوذة من مبادرات متخذة على مستوى الدولة وعلى المستوى الدولي لبناء سرديات مضادة للمعلومات المُضلّلة توفّر معلومات واقعية تهدف إلى دحض الأكاذيب المترسخة في السرديات المُضلّلة. كما ويناقش ما إذا كان دحض المعلومات المُضلّلة هو استجابة فعالة للتضليل استناداً إلى أحدث الدراسات العلمية حول هذا الموضوع. ومتى كانت الحملات الحكومية لمُكافحة التّضليل شفافةً وتهدف إلى تعزيز الحوار مع المواطنين؛ بمقدورها أن تُعزّز الثقة في السُّلطات وشفافيتها. والخطر الكامن، مع ذلك، هو أن هذه الآليات تشكل مبادرات تواصل استراتيجي أُحاديّ الاتجاه تخدم مصالح السياسيين ولا تعالج بعض الأسباب الكامنة وراء التّضليل التي تستلزم سياسات تفوق المستوى المعلوماتي (مثل: التنمية الاقتصادية للفئات والمناطق المهمشة). بالتالي، يمكن لبعض مبادرات مُكافحة التّضليل أن تُعمّق الانقسامات الحزبية.

1.2.5 ما هي الأمور ومن هي الجهات التي ترصدها أو تستهدفها هذه الاستجابات؟

إنّ مبادرات مُكافحة التّضليل التي تطلقها السُّلطات الوطنية والدولية تستهدف حملات التّضليل الخارجية والداخلية على حد سواء. وفي حين أنّ بعض المبادرات لا تستهدف نوعاً محدداً من المعلومات المُضلّلة يركّز البعض الآخر على أمور معيّنة، مثل فريق التواصل الاستراتيجي للشرق التابع لدائرة العمل الخارجي الأوروبية الذي يركّز في المقام الأول على رصد التّضليل النابع من دول خارج الاتحاد الأوروبي. ويتم بلورة بعض المبادرات لكشف زيف التّضليل خلال الفترات الانتخابية، مثل: الموقع الذي أقامته المحكمة الانتخابية العليا في البرازيل في الفترة التي سبقت الانتخابات العامة عام 2018. والمعلومات المُضلّلة المتعلقة بالمسائل الصحية هي أيضاً مصدر قلق أدى إلى إطلاق العديد من المبادرات المخصصة لمُكافحة التّضليل، خاصةً في سياق جائحة فيروس كورونا.

2.2.5 مَن هي الجهات التي تسعى هذه الاستجابات إلى مُساعدتها؟

تُركّز العديد من هذه الحملات والمبادرات على إعلام الجمهور العام بالإدعاءات المُضلّلة التي يتم اكتشافها، كما هي الحال في السياقات الانتخابية، حول مجموعة من السياسات، والكوارث الطبيعية، ومسائل الصحة العامة ومخاوف تتعلق بالسلامة وغيرها. بالإضافة إلى ذلك، فإنّ الامتداد الدولي لحملات مُكافحة التّضليل قد يُصمم أيضاً بطريقة تحافظ وتحسّن من نظرة المجتمع الدولي للبلد ولحكومته (أو التكتل الإقليمي). وتتراوح هذه المبادرات من الدبلوماسية العامة إلى الدعاية. وجزء من هذه المبادرات يزوّد الجهات العسكرية بالتحليلات اللازمة، كالعمل الذي يقوم به «مركز التميّز للتواصل الاستراتيجي التابع لحلف شمال الأطلسي»، الذي يتولى نشر التقارير ودعم قدرات التواصل الاستراتيجي للحِلف.

وتستند حملات مُكافحة التّضليل على دحض المعلومات المُضلّلة، وتفترض تلك الدول التي تطلق حملات مُكافحة التّضليل أن كشف المعلومات المُضلّلة وتزويد الجمهور بمعلومات واقعية صحيحة من شأنهما الحد من تصديق المعلومات غير الواقعية والحد من أثرها. كذلك، تسعى هذه الدول إلى إثارة شكوك الجمهور إزاء هذه المعلومات بناءً على مَنشأها. وتذهب بعض المبادرات أبعد من مسألة تقديم معلومات واقعية، فتعمد إلى طرح السرديات والوقائع من زاويةٍ مختلفة مرتكزةً على نفوذها الجيوسياسي.

3.2.5 ما هي مخرجات الحملات الوطنية والدولية لمُكافحة التّضليل؟

ترتكز مبادرات مُكافحة التّضليل على أنشطة التحقُّق من الوقائع ونشر معلومات تُعتبر رسمياً؛ معلومات موثوقة. وتوضَع عملية التحقُّق على الإنترنت ويصار إلى مشاركتها على وسائل التواصل الاجتماعي في محاولة للوصول إلى الجمهور عبر المنصّات نفسها التي تُنشر عليها المعلومات المُضلّلة. ويمكن أيضاً أن توضع عملية كشف زيف المعلومات المُضلّلة مباشرةً على قنوات التواصل الاجتماعي، على غِرار الحساب الباكستاني Pakistani@FakeNews_Buster. ويمكن أيضاً أن تُعرض أعمال الرصد في تقارير وتحليلات واسعة النطاق تغذي جهود التواصل الاستراتيجي، على غرار التقارير التي يضعها «مركز التميّز للتواصل الاستراتيجي التابع لحلف شمال الأطلسي» و«فريق الشرق للتواصل الاستراتيجي التابع لدائرة العمل الخارجي الأوروبية». بالإضافة إلى ذلك، قد تتم مشاركة التقارير مع وسائل الإعلام لتأمين التغطية الإعلامية.

4.2.5 مَن هي الجهات الفاعلة الرئيسية ومَن هي الجهات التي تُموّل هذه الاستجابات؟

إنّ المبادرات الواردة أدناه، التي تم جمعها من خلال البحوث التي تمت حتى شهر آذار/مايو من عام 2020، منبثقة عن حكومات أو منظمات دولية، وبالتالي فهي ممولة من السُّلطات.

حملات مُكافحة التّضليل	الجهات الفاعلة: رابطة أُمم جنوب شرق آسيا، البرازيل، كمبوديا، كندا، الصين، جمهورية الكونغو الديمقراطية، والاتحاد الأوروبي/دائرة النفاذ الخارجي الأوروبية، الهند، إندونيسيا، ماليزيا، المكسيك، حلف شمال الأطلسي، عُمان، باكستان، الاتحاد الروسي، جنوب أفريقيا، تايلاند، تونس، المملكة المتحدة والأمم المتحدة واليونسكو ومنظمة الصحة العالمية	
	1	المحكمة الانتخابية العليا في البرازيل
	2	برنامج تلفزيون كمبوديا
	3	أنشطة كندا في إطار مبادرة المواطن الرقمي
	4	المنصة الحكومية الصينية Piyao لدحض الشائعات
	5	استجابة جمهورية الكونغو الديمقراطية للمعلومات المغلوطة والمُضلّلة حول وباء إيبولا
	6	فريق التواصل الاستراتيجي للشرق التابع لدائرة العمل الخارجي الأوروبية
	7	فرع حرب المعلومات في الجيش الهندي
	8	وحدة التحقُّق من الوقائع في وزارة الإعلام والإذاعة في الهند
	9	بوابة CEKHOAKS! لكشف زيف المعلومات المُضلّلة في إندونيسيا

	10	بوابة Sebenarnya.my لكشف زيف المعلومات المُضلّلة في ماليزيا
	11	موقع Verificado لوكالة الأنباء المكسيكية Notimex
	12	مركز التميز للتواصل الاستراتيجي التابع لحلف شمال الأطلسي
	13	مبادرة التواصل الصادرة عن الحكومة العُمانية
	14	حساب FakeNews_Buster على تويتر لدولة باكستان
	15	صفحة كشف زيف المعلومات المُضلّلة التابعة لوزارة الشؤون الاجتماعية في الاتحاد الروسي
	16	مركز تايلاند لمُكافحة الأخبار الزائفة
	17	موقع Tunisia's Check News
	18	برنامج وزارة الخارجية والكومنولث لمُكافحة التّضليل وتطوير الإعلام
	19	حملة منظمة الصحة العالمية لدحض الخرافات حول جائحة فيروس كورونا Coronarvirus Mythbusters*
	20	استجابة الأمم المتحدة للتواصل (حول جائحة فيروس كورونا*
	21	حملات اليونسكو لمُكافحة التّضليل حول جائحة فيروس كورونا *
	22	شراكة رابطة أمم جنوب شرق آسيا لمُكافحة التّضليل المتعلق بجائحة فيروس كورونا*
	23	حملة الاتحاد الأوروبي لدحض الخرافات حول فيروس كورونا mythbusting *
	24	الصفحة الإلكترونية لحملة جنوب أفريقيا الخاصة بجائحة فيروس كورونا
	25	حملة الهند على تطبيق واتساب لمُكافحة التّضليل حول جائحة فيروس كورونا
	26	وحدة الاستجابة السريعة في حكومة المملكة المتحدة لمُكافحة التّضليل حول فيروس كورونا

* ترد تفاصيل هذه المبادرات في دراسة الحالة أدناه حول فيروس كورونا

الجدول 4. الحملات الوطنية والدولية لمُكافحة التَّضليل

1. المحكمة الانتخابية العليا في البرازيل (2018)

أطلقت المحكمة الانتخابية العليا في البرازيل موقعاً خاصاً بها للتحقق من الوقائع ومُكافحة التّضليل[129] (Brazil Superior Electoral Court, 2018) في الفترة التي سبقت الانتخابات العامة في اكتوبر/تشرين الأول من عام 2018. وقد أحالت تقارير التّضليل التي وردتها إلى مكتب المُدّعي العام والشرطة الاتحادية للتحقق منها.

2. برنامج تلفزيون كمبوديا (2019)

في أوائل عام 2019 أطلقت وزارة الإعلام في كمبوديا برنامجاً تلفزيونياً أسبوعياً مباشراً يبث على التلفزيون الوطني لمُكافحة المعلومات التي تعتبرها مضللةً (Dara, 2019).

....

129 http://www.tse.jus.br/hotsites/esclarecimentos-informacoes-falsas-eleicoes-2018/

3. برنامج أنشطة كندا في إطار مبادرة المواطن الرقمي (2019)

في 2019، قامت كندا بتمويل سلسلة من المبادرات التي تهدف إلى رفع مستوى الوعي حول مشكلة التّضليل وبناء القدرة على مُكافحة هذه المشكلة في صفوف الجمهور العريض (Canada Government, 2019c).

4. المنصة الحكومية الصينية Piyao لدحض الشائعات (2018)

أطلقت الحكومة الصينية منصة Piyao (أي «دحض الشائعات»)[130] التي تستضيفها لجنة الشؤون السيبرانية المركزية بالتعاون مع «وكالة أنباء شينخوا الرسمية»، في آب/أغسطس 2018. تشجع المنصة المواطنين على الإبلاغ عن المعلومات المُضلّلة وتستخدم الذكاء الاصطناعي للتنقيب عن الشائعات. كذلك، تنشر الأخبار التي تحظى بموافقة الدولة والمعلومات المضادة للتضليل. وتعمل المنصة كمكان يجمع جهود الوكالات الحكومية الصينية لدحض ما تعتبره معلومات مضللة (Qiu & Woo, 2018).

5. استجابة جمهورية الكونغو الديمقراطية للمعلومات المغلوطة والمُضلّلة حول مرض فيروس إيبولا (2018)

رداً على انتشار الشائعات والمعلومات المغلوطة والمُضلّلة عن «فيروس إيبولا» في جمهورية الكونغو الديمقراطية، قامت المنظمات الصحية بالتعاون (أي اليونيسف ومنظمة الصحة العالمية والاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر) لإنشاء قاعدة بيانات تضم الشائعات التي تنتشر داخل المجتمعات وعبر القنوات الإعلامية والاجتماعية. ولأنّ التّضليل قد يُصعّب عمل الطاقم الطبي على الميدان، قدمت منظمة الصحة العالمية المشورة في ما يخص التحقُّق في الوقائع والتواصل حول المخاطر، وذلك للمتطوعين والعاملين في الخطوط الأمامية لمُكافحة فيروس إيبولا (WHO, 2018). كذلك، قامت وزارة الصحة بتعبئة المواطنين للإبلاغ عن المعلومات المُضلّلة والمغلوطة التي تنتشر على تطبيق واتساب. وتهدف جهود المتابعة والرصد هذه إلى تطوير الاستراتيجية الأكثر ملائمةً للاستجابة للتضليل ولدحض المعلومات المُضلّلة التي تُنقل من شخص إلى شخص أو عبر الإذاعة وعبر تطبيق واتساب (Spinney, 2019; Fidler, 2019).

6. فريق التواصل الاستراتيجي للشرق التابع لدائرة العمل الخارجي الأوروبية (2015)

في آذار/مارس 2015، كُلّف المجلس الأوروبي الممثل الأعلى بالتعاون مع مؤسسات الاتحاد الأوروبي والدول الأعضاء بتقديم خطة عمل حول التواصل الاستراتيجي. وكجزء من العمل الهادف إلى تحسين التوقع والتصدي والاستجابة لأنشطة التّضليل التي تقوم بها جهات خارجية؛ تمَّ إنشاء فريق العمل كجزء من دائرة العمل الخارجي الأوروبي للتصدي لحملات التّضليل الخارجية.

ولهذا الهدف، تم إنشاء فريق صغير داخل الدائرة لتطوير رسائل إيجابية عن الاتحاد الأوروبي في دول الجوار الشرقي. وقد كُلّف أيضاً بدعم البيئة الإعلامية في هذه المنطقة. وأخيراً، قام الفريق بتحليل اتجاهات التّضليل وكشف السرديات المُضلّلة النابعة من مصادر خارج الاتحاد الأوروبي. ويمكن الاطّلاع

[130] http://www.piyao.org.cn

على عمل الفريق في مجال التصدي للتضليل على موقعه على الإنترنت[131] (euvsdisinfo.eu). وللفريق موقع باللُّغة الروسية (eeas.europa.eu/ru/eu information-russian_ru)[132] يهدف إلى نشر المعلومات حول أنشطة الاتحاد الأوروبي في دول الجوار الشرقي (EU EEAS, 2018).

كذلك، لفريق التواصل الاستراتيجي «نظام إنذار مُبكّر» بين الدول الأعضاء في الاتحاد الأوروبي، أُطلق في آذار/مارس 2019 كعنصر من مدوّنة الممارسة الأوروبية حول التّضليل. وقد استُخدم للمرة الأولى في سياق جائحة فيروس كورونا (Stolton, 2020)

7. فرع حرب المعلومات في الجيش الهندي (2019)

وافقت وزارة الدفاع الهندية على إنشاء فرع مختص بحرب المعلومات داخل الجيش لمواجهة ما تعتبره تضليلاً ودعايةً وذلك في آذار/مارس 2019 (Karanbir Gurung, 2019).

8. وحدة التحقُّق من الوقائع في وزارة الإعلام والإذاعة في الهند (2019)

في وقت لاحق، في تشرين الثاني/نوفمبر 2019، أعلنت الحكومة الهندية عن إنشاء وحدة للتحقق من الوقائع في وزارة الإعلام والإذاعة. وستعمل هذه الوحدة وفقاً «لمبادئ أربعة وهي: البحث عن المعلومات وتقييمها ووضع الاستجابات والاستهداف المعروفة باللُّغة الإنكليزية بمختصر (FACT)». كما وستبلّغ عن التّضليل للوزارات الحكومية المعنيّة (Mathur, 2019).

9. بوابة CEKHOAKS! لكشف زيف المعلومات المُضلّلة في إندونيسيا

تسمح بوابة CEKHOAKS![133] لكشف زيف المعلومات المُضلّلة في إندونيسيا للمواطنين بالتبليغ عن المعلومات المُضلّلة والخدع، والاطلاع على المنشورات التي تم كشف زيفها. ويدعم هذا الموقع الوزارة الإندونيسية للاتصالات السلكية واللاسلكية، وجمعية مُكافحة التشهير الإندونيسية، وجهات حكومية أخرى ومنظمات من المجتمع المدني.

10. بوابة Sebenarnya.my لكشف زيف المعلومات المُضلّلة في ماليزيا (-2017)

قامت لجنة الاتصالات والوسائط المتعددة في ماليزيا بإنشاء بوابة لكشف المعلومات المُضلّلة؛ تحت مُسمّى «sebenarnya.my»[134] في آذار/مارس 2017 ووضعت تطبيقاً مرافقاً لها في آذار/مارس 2018 لرفع الوعي والحد من انتشار التّضليل على الإنترنت (Buchanan, 2019).

....

[131] http://euvsdisinfo.eu/

[132] https://eeas.europa.eu/ru/eu-information-russian_ru

[133] https://stophoax.id

[134] https://sebenarnya.my/

11. موقع Verificado لوكالة الأنباء المكسيكية (-2019) Notimex

في حزيران/يونيو 2019، أطلقت وكالة الأنباء المكسيكية موقعها الخاص بالتحقُّق من الوقائع ومُكافحة التّضليل وهو «Verificado NTX».[135]

12. مركز التميّز للتواصل الاستراتيجي التابع لحلف شمالي الأطلسي (2014)

إنّ «مركز التميّز للتواصل الاستراتيجي التابع لحلف شمالي الأطلسي» ومقره في ريغا، لاتفيا، هو مركز معتمد من منظمة حلف شمال الأطلسي، تم إنشاؤه عام 2014 بموجب مذكرة تفاهم بين إستونيا وألمانيا وإيطاليا ولاتفيا وليتوانيا وبولندا والمملكة المتحدة. وهو مستقل عن هيكل قيادة حلف شمال الأطلسي ولا ينطق باسم الحلف. وقد انضمت هولندا وفنلندا إلى المركز عام 2016، والسويد عام 2017، وكندا في 2018 وسلوفاكيا في أوائل 2019. وكان من المفترض أن تنضم إليه فرنسا والدنمارك في عام 2020. ويعمل المركز على تحليل المعلومات المُضلّلة وتقديم الدعم لقدرات التواصل الاستراتيجي لحلف شمال الأطلسي (NATO Stratcom COE, 2019).

13. مبادرة التواصل الصادرة عن الحكومة العُمانية (2018)

قدم المركز العُماني للتواصل الحكومي تدريباً لمساعدة دوائر الإعلام والتواصل داخل المؤسسات الحكومية على رصد ودحض المعلومات المُضلّلة (Al Busaidi, 2019).

14. حساب FakeNews_Buster على تويتر لدولة باكستان (2018)

أطلقت وزارة الإعلام والإذاعة في باكستان حساباً على تويتر [136] (FakeNews_Buster) لرفع مستوى الوعي ودحض المعلومات المُضلّلة في تشرين الأول/أكتوبر 2018 (Dawn, 2018). ومن التغريدات المتكررة «إنّ نشر الأخبار الزائفة FakeNews# ليس فقط عملاً غير أخلاقي وغير قانوني ولكنه أيضاً إساءة للأُمّة.» وتقع على عاتق الجميع مسؤولية رفض هذا السلوك غير المسؤول. ”Reject #FakeNews (@FakeNews_Buster).

15. صفحة كشف زيف المعلومات المُضلّلة التابعة لوزارة الشؤون الاجتماعية في الاتحاد الروسي (2017)

لوزارة الشؤون الخارجية في الاتحاد الروسي صفحة على الإنترنت مخصصة لرفع مستوى الوعي حول المواد المنشورة التي تتضمن معلومات عن الاتحاد الروسي يعتبرها الاتحاد خاطئة.[137] وبعد التعديلات التي أُدخلت على قانون المعلومات عام 2019، كان من المتوقع أن تقوم الجهة الناظمة للإعلام في الاتحاد الروسي Roskomnadzor بإنشاء «قاعدة بيانات حول الأخبار الزائفة» (Zharov, 2019).

135 http://verificado.notimex.gob.mx

136 https://twitter.com/FakeNews_Buster

137 https://www.mid.ru/en/nedostovernie-publikacii

16. مركز تايلاند لمُكافحة الأخبار الزائفة (2019)

أنشأ وزير الاقتصاد والمجتمع الرقميين في تايلاند مركزاً دولياً «لمُكافحة الأخبار الزائفة في تشرين الأول/أكتوبر 2019 لرصد المعلومات المُضلّلة ودحضها، وتُعرّف على أنها: «أيّ مُحتوى واسع الانتشار على الإنترنت يُضلّل الأفراد أو يُلحق الضرر بصورة البلد» (Tanakasempipat, 2019b). وبالتنسيق مع السُّلطات المختصة، يتم نشر إشعارات التصحيح على حسابات المركز على وسائل التواصل الاجتماعي، وعلى موقعه الإلكتروني [138] (antifakenewscenter.com) وفي الصحف. وقد أصدر المركز أيضاً مذكرات استيقاف وقبض (Bangkok Post, 2019).

17. موقع Tunisia's Check News (2019-)

أُطلق هذا الموقع التونسي للتحقق من الوقائع وكشف زيف المعلومات المُضلّلة (tunisiachecknews.com) [139] في تشرين الأول/أكتوبر 2019 وتشرف عليه الهيئة التونسية العليا المستقلة للاتصال السمعي البصري (HAICA) التي تعمل بتعاون وثيق مع صحافيين من المؤسسات الإعلامية الحكومية (التلفزيون الوطني والإذاعة الوطنية ووكالة تونس إفريقيا للأنباء).

18. برنامج وزارة الخارجية والكومنولث لمُكافحة التّضليل وتطوير الإعلام (2016–2021)

في نيسان/أبريل 2018، وفي سياق التّضليل حول حادث التسمم في «ساليسبري» (Symonds, 2018)، أطلقت وزارة الخارجية والكومنولث في المملكة المتحدة بالتعاون مع وزارة الدفاع ووزارة الثقافة والإعلام والرياضة ومكتب رئيس الوزراء، برنامجاً «لمُكافحة التّضليل ولتطوير الإعلام». هذا المشروع هو جزء من مجموعة «برامج صندوق النزاعات والاستقرار والأمن في أوروبا الشرقية وآسيا الوسطى ومنطقة البلقان الغربية».

ويوفر البرنامج الدعم المالي والتوجيه للمنظمات بهدف «تحسين نوعية الخدمة العامة ووسائل الإعلام المستقلة (بما في ذلك باللغة الروسية) لتمكينها من دعم التماسك الاجتماعي، وترسيخ القيم العالمية وتزويد المجتمعات المحلية في بلدان أوروبا الشرقية بإمكانية النفاذ إلى معلومات موثوقة». ومن خلال دعم جهود المجتمع المدني لكشف التّضليل، يتوقع البرنامج تعزيز صمود المجتمع في أوروبا.[140]

5.2.5 دراسة حالة بشأن الاستجابة: التّضليل المتعلق بجائحة فيروس كورونا

لقد كانت حملات مُكافحة التّضليل عناصر قوية من الاستجابات المعتمدة على المستوى الوطني والدولي للتصدي للمعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا. وقد تم تنفيذها بسرعة بهدف تعبئة المجتمعات الإلكترونية ونشر المعلومات الصحيحة الرسمية وكشف زيف المُحتوى غير الصحيح. وقد عُقدت الشراكات بين مختلف شركات وسلطات التواصل على الإنترنت لتوفير قنوات تفاعلية لنشر المُحتوى الرسمي. وتشمل التدابير في هذه الفئة الحملات والوحدات الخاصة التي أُنشِأت بهدف إنتاج محتويات لمُكافحة التّضليل.

138 https://www.antifakenewscenter.com/

139 https://tunisiachecknews.com/

140 إنّ «مختبر الاتحاد الأوروبي للمعلومات المُضلّلة» الذي قام بصياغة هذا القسم 5.2 هو حائز على منحة من برنامج وزارة الخارجية والكومنولث لمُكافحة التّضليل ولتطوير الإعلام.

ومـن الأمثلـة علـى هـذا النـوع مـن الاسـتجابات التـي طُـوّرت لمُكافحـة التّضليـل المتعلـق بجائحـة فيـروس كورونـا الاسـتجابات التاليـة:

- صفحة منظمة الصحة العالمية لدحض الخرافات mythbusting: في مؤتمر صحفي، أعلن مسؤول في منظمة الصحة العالمية «أننا بحاجة إلى لقاح ضد التّضليل» (WHO, 2020). بعد انتشار جائحة فيروس كورونا، أنشأت منظمة الصحة العالمية صفحةً رسميةً[141] لتصحيح الخرافات؛ تسعى إلى توفير معلومات موثوقة عن هذا الفيروس وتبعاته الخطيرة. كذلك أنشأت موقعاً لشبكة المعلومات حول الأوبئة التابعة لمنظمة الصحة العالمية أو EPI-WIN.[142] (هذه المبادرة هي رقم 19 في الجدول أعلاه)

- أطلق الأمين العام للأمم المتحدة استجابة الأمم المتحدة في مجال التواصل «للإغداق على الإنترنت بالحقائق والعلم»، في سياق مُكافحة آفة التّضليل المتنامية التي وصفها بأنها «السُّم الذي يهدد حياة المزيد من الأشخاص» (UN News, 2020; UN Department of Global Communications, 2020). وفي آيار/مايو، تم تنفيذ هذه المبادرة التي سُمّيت "Verified" (بما معناه: تمَّ التحقُّق منه) وهدفها إنشاء طاقم من «المستجيبين الرقميين الأوائل» يعمل على تعزيز حجم وتوسيع نطاق وصول المعلومات الموثوق بها والدقيقة المحيطة بالجائحة.[143] (هذه المبادرة هي رقم 20 في الجدول أعلاه)

- نشرت اليونسكو إيجازين سياسيين لفهم «أزمة المعلومات المُضلّلة» (Posetti & Bontcheva, 2020a; Posetti & Bontcheva, 2020b) شكّلا جزءً من حملة أوسع نطاقاً تهدف إلى مُكافحة التّضليل والتأثير على عملية وضع السياسات على مستوى الدولة. كما وأنتجت منشورات باللُّغات المحلية تحت عنوان «تمزيق المعلومات المُضلّلة» أو shredder misinformation.[144] (هذه المبادرة هي رقم 21 في الجدول أعلاه). كما وأطلقت اليونسكو حملةً عالميةً سُمّيت FACTS أي «حقائق» خلال الاحتفال باليوم العالمي لحُريّة الصحافة في 3 أيار/مايو من عام 2020، وتم إنتاج محتويات مسموعة بلغات عدة لإذاعتها على محطات الإذاعة في جميع أنحاء العالم. ثم أُطلقت في وقت لاحق مبادرة أخرى بعنوان «لا للانتشار الواسع» أو Don't Go Viral.[145]

- تعمل فرق «مبادرة النبض العالمي» للأمم المتحدة في نيويورك وكمبالا وإندونيسيا على رفع الوعي الظرفي حول تفشي وظهور وانتشار «أزمة المعلومات» التي من شأنها أن تدعم جهود كافة أركان الأمم المتحدة. كما وتعمل على إجراء تحليلات هدفها تحديد الجهود التي تُعزّز انتشار وتأثير المعلومات الصحية[146]. وتحقيقاً لهذه الغاية، تعمل هذه الفرق على إنشاء وتوسيع نطاق الأدوات

141 https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters

142 https://www.epi-win.com/advice-and-information/myth-busters

143 https://www.shareverified.com/en.
https://news.un.org/en/story/2020/05/1064622
https://www.shareverified.com/en
https://news.un.org/en/story/2020/05/1064622

144 https://en.unesco.org/news/faq-covid-19-and-misinformation-shredder-african-local-languages

145 https://en.unesco.org/commemorations/worldpressfreedomday/facts-campaign;
https://en.unesco.org/covid19/communicationinformationresponse/dontgoviral
https://en.unesco.org/covid19/communicationinformationresponse/audioresources

146 https://www.unglobalpulse.org/project/understanding-the-covid-19-pandemic-in-real-time/

والمنهجيات والأطر التحليلية لدعم كيانات الأمم المتحدة وتحسين فهمها للسياقات التشغيلية التي تتم فيها مُكافحة الآثار السلبية لفيروس كورونا في أفريقيا. وبناءً على منهجيات علمية، يركّز الدعم المباشر المقدم لمنظمة الصحة العالمية في أفريقيا على تقديم الدعم والمنتجات التحليلية التي تعتمد على المنهجيات التالية: 1) تحليل كمي ونوعي قصير الأجل للإشارات الرقمية بناءً على الشائعات والمعلومات المغلوطة التي تبلّغ عنها المكاتب الميدانية. 2) المراقبة المستمرة على أساس تصنيف متكيّف يسمح بتحديد «أزمات المعلومات المُضلّلة» سريعة التطوّر والتقييم الكمي للتطوّر الزمني لموضوعاتٍ معينة. وهذا يشمل تحليلات تنبؤية لشائعات ومفاهيم وفقاً للحجم والامتداد الجغرافي والقنوات المستخدمة؛ 3) تحليل للمشاعر والعواطف حول مفاهيم معينة، بما في ذلك ظهور وتصاعد خطاب الكراهية. هذا سوف يسمح للفرق بتطوير إطار يسعى إلى تحقيق الاستفادة القصوى من المعلومات التي توفرها منظمة الصحة العالمية وشركاؤها لمواجهة التّضليل.

- اجتمع وزراء خارجية «رابطة أمم جنوب شرق آسيا» وجمهورية الصين الشعبية لتنسيق العمل ضد فيروس كورونا. وعلى وجه الخصوص، اتفق الوزراء على تعزيز تعاونهم في مجال التواصل المتعلق بالمخاطر «لضمان وصول المعلومات الصحيحة والدقيقة حول فيروس كورونا لكافة الناس ولتفادي انسياقهم وراء المعلومات المُضلّلة والأخبار الزائفة المتعلقة بالفيروس» (ASEAN, 2020). على الرغم من ذلك، لم يتم تفصيل كيفية تطبيق هذا التعاون على أرض الواقع. (هذه المبادرة هي رقم 22 في الجدول أعلاه)

- نشر البرلمان الأوروبي توجيهات تتعلق بالتعامل مع الخرافات حول جائحة فيروس كورونا.[147] (هذه المبادرة هي رقم 23 في الجدول أعلاه)

- أصدرت حكومة جنوب أفريقيا قراراً يلزم كافة المواقع الإلكترونية العاملة على نطاق zaDNA بأن تتضمن صفحتها الأساسية رابطاً واضحاً لموقع www.sacoronavirus.co.za (وهو الموقع الوطني حول جائحة فيروس كورونا).[148] (هذه المبادرة هي رقم 24 في الجدول أعلاه)

- أطلقت الحكومة الهندية «روبوت دردشة» على تطبيق واتساب، يهدف إلى مُكافحة التّضليل حول جائحة فيروس كورونا (Chaturvedi 2020). (هذه المبادرة هي رقم 25 في الجدول أعلاه

- أنشأت وزارة الشؤون الرقمية والثقافة والإعلام والرياضة في المملكة المتحدة وحدةً مخصصةً لرصد التّضليل حول الجائحة والاستجابة له؛ تعمل بشكل منتظم بالتعاون مع شركات التواصل عبر الإنترنت (Sabbagh, 2020). وشملت هذه المبادرة «وحدةً للاستجابة السريعة» مصممةً بهدف «الحد من انتشار الأكاذيب والشائعات التي قد تهدد الأرواح» (UK Parliament, Sub Committee on Online Harms and Disinformation, 2020). واستكمالاً لهذا الجهد، قامت وزارة التنمية الدولية بدعم مبادرة تهدف إلى الحد من انتشار التّضليل المتعلق بهذا الوباء، لا سيما في جنوب شرق آسيا وأفريقيا. وقد ركّز البرنامج على التحقّق من المعلومات بمساعدة من المؤسسات الإعلامية الشريكة على غرار BBC. كما وركّز على مشاركة الأخبار الموثوق بها بمساعدة من مؤثرين مختارين (UK Department for International Development, 2020). (هذه المبادرة هي رقم 26 في الجدول أعلاه).

---

147 https://www.europarl.europa.eu/news/en/headlines/society/20200326STO75917/disinformationhow-to-recognise-and-tackle-covid-19-myths

148 https://sacoronavirus.co.za

6.2.5 كيف يتم تقييم الاستجابات الوطنية والدولية لمُكافحة التّضليل؟

بما أنّ العديد من المُبادرات المعروضة في هذا القسم حديثة للغاية، ليس هناك من أدلة كثيرة تسمح بإجراء تقييم ذو مغزى. وفي الوقت نفسه، يبدو أن المبادرات لم تتضمن بشكل صريح في خططها أنشطة متابعة ورصد لتقييم أثرها وفعاليتها المقصودين (وغير المقصودين). وبما أنّ هذه المبادرات ممولة من الحكومات، يُفترض أن يتم تقييم فعاليتها داخلياً من قبل الأجهزة الحكومية، أو خارجياً من قبل المجتمع المدني والمؤسسات الإعلامية التي تسعى إلى تعزيز المُساءلة والشفافية. وفي سياق العمليات الدولية، تقوم الدول الأعضاء بتقييم المبادرات التي تدعمها. على سبيل المثال، في نيسان/أبريل 2018 قام مجلس الشؤون الخارجية في الاتحاد الأوروبي (EU Foreign Affairs Council, 2018) بالثناء على «العمل الذي يقوم به فريق التواصل الاستراتيجي للشرق» في سياق تعزيز قدرة الاتحاد الأوروبي وجيرانه على الصمود؛ وهو ما يعتبره المجلس حاجةً ضروريةً.

7.2.5 التحديات والفرص

يشعر الجمهور أنّ حملات مُكافحة التّضليل شرعية ومقنعة متى كانت المؤسسات المسؤولة عنها موثوق بها. ويمكن أن تبقى استراتيجيات كشف الزيف داخل حدود حُريّة التّعبير، عن طريق دحض المُحتوى غير المحظور لكونه «غير صحيح». وعندما تُبنى هذه الحملات على الوقائع وتخضع للتدقيق، من المفترض أن تكون أكثر فعاليةً من الجهود السرية و/أو تلك التي تُبنى على السرديات لدرجة تجعلها أقرب إلى الدعاية.

وقد ترفض الجماهير النقدية وغير المعنية عملية دحض «المعلومات المُضلّلة» حيث تعتبرها مبادرات علاقات عامة تعتمدها الجهات الحكومية، لا عمليةً محايدةً للتحقق من الوقائع. وهذا من شأنه أن يغذي الشكوك ونظريات المؤامرة حول تدخل الدولة، كما ومن شأنه أن يرسخ عدم الثقة في الجهات الحكومية، خاصةً تلك التي لديها تاريخ من الرقابة والدعاية. ويُضاف إلى ذلك خطر قيام الحكومات بإصدار نسختها من «الحقائق البديلة» كخطوة لزرع بذور التّضليل. ومتى كانت الجهات الفاعلة تعتمد هي أيضاً أساليب التّضليل، قد يكون ذلك من العوامل التي تؤدي إلى ارتداد عملها في مجال كشف الأكاذيب.

ومن النظريات المعروفة في مجال التواصل، نظرية «أثر باربرا سترايسند» ومفادها أنّ أيّ محاولة لإخفاء أو فرض رقابة على معلومة معيّنة قد ترتد لتؤدي إلى نتيجة عكسية غير مقصودة، وهي أن تنتشر هذه المعلومة بشكل واسع في هذا العصر الرقمي (Masnick, 2003). وقد سُميت هذه النظرية على اسم المغنية «باربرا سترايسند» التي حاولت إزالة صورة جوية لمنزلها في «ماليبو» فكان لمحاولتها أثر عكسي فجذبت مزيداً من الانتباه للمنزل. ويمكن اختبار هذه الفرضية في ما يتعلق بكشف التّضليل أيضاً، بما في ذلك في حال المبادرات الحكومية.

في العلوم المعرفية، يُعرف هذا الأثر باسم «النتيجة العكسية»، ومعناه أن دحض المعلومات قد يؤدي إلى ترسيخها في عقل القارئ (Cook et al., 2014; Nyhan & Reifler, 2006). وفي تحليل الفعالية النفسية لرسائل مُكافحة التّضليل، يُوصي بعض الباحثين أنه عندما يكون هناك حاجة لتكرار كذبة ما بهدف كشف زيفها، فمن الأفضل الحد من تكرارها (Sally Chan et al., 2017).
لكن البحوث الحديثة لم تجد أيّ دليل يثبت أن الدحض الذي يكرر الادعاءات الكاذبة ويحددها على أنها كذلك يؤدي إلى زيادة الإيمان بهذه المعلومات المُضلّلة. على العكس، تبيّن أن تقديم شرح مفصل بديل

لمعلومات ما هو أكثر فعاليةً (Ecker et al., 2019). وتشير بعض الأبحاث إلى أنه لابد لعملية كشف الزيف أن تستخدم طريقة «ساندوتش الحقيقة» كما وصفها عالم اللُّغة «جورج لاكوف» (Hanly, 2018)، حيث تقوم المعلومات الصحيحة بلف المعلومات الخاطئة من جانبيها، فلا تُعطى المعلومات الخاطئة لا المكانة الأولى ولا المكانة الأخيرة في السردية.[149] ومع ذلك، لابد من مزيد من البحث في الفروقات بين عملية كشف الزيف التي تقوم بها الحكومات وتلك التي تقوم بها جهات مستقلة.

ومن حدود الدحض أنه يعمل فقط على الادعاءات الكاذبة التي تم اكتشافها. ويأخذ التّضليل أشكالاً مختلفةً لا تتضمن بالضرورة ادعاءات كاذبة بشكل صريح، حيث قد يتضمن استخدام معلومات معينة مأخوذة خارج سياقها أو تطبيق معلومات صحيحة بشكل مُضلّل. وغالباً ما يصار إلى التّعبير عن هذه المعلومات بطرق لها صدى عاطفي قوي.

في ما يتعلق بالفرص، يمكن للحملات التي تقودها السُّلطات العامة حشد موارد كبيرة –مالية وبشرية– لرصد المُحتوى والتحقُّق منه وتعميم النتائج. وإنّ الطابع العام لهذه المبادرات قد يؤدي أيضاً إلى تفاعل ونقاش في المؤسسات العامة من خلال البرلمان وغيره من آليات الرقابة.

8.2.5 توصيات للحملات الوطنية والدولية لمُكافحة التّضليل

توصيات للدول:

- العمل بشكل أوثق مع منظمات المجتمع المدني والوكالات الإخبارية والخبراء الأكاديميين لبلورة حملات مطلعة تهدف إلى الاستجابة لأنواع مختلفة من التّضليل.
- النظر في تنظيم حملات تهدف إلى رفع مستوى الوعي حول قيمة الصحافة والصحفيين المستقلين والنقديين ودورهم في حماية المجتمعات من التّضليل.
- الاستثمار في البحوث التي تقيس فعالية حملات مُكافحة التّضليل.

توصيات للباحثين:

- إجراء بحوث حول الجمهور لاختبار استجاباته للحملات الوطنية والدولية المختلفة (مثلاً الحملات على الإنترنت/خارج الإنترنت، والحملات التفاعلية والسمعية والبصرية) وذلك بين مجموعات مختلفة (كالأطفال والشباب وكبار السن والمجتمعات من خلفيات اجتماعية واقتصادية مختلفة، وأصحاب المعتقدات السياسية المتنوعة، والأشخاص الذين تم تحديدهم على أنهم عُرضة للتأثر بالمعلومات المُضلّلة و/أو مشاركتها).

توصيات لشركات التواصل عبر الإنترنت:

- زيادة الدعم المالي وتحسين بروز الحملات الحكومية لمُكافحة التّضليل إلى مرحلة ما بعد الأزمات على غِرار جائحة فيروس كورونا.

[149] يُرجَى مراجعة المناقشة المفصلة للأدبيات في الفصل الثالث.

3.5 الاستجابات الانتخابية

المؤلفون: دينيس تيسو، جولي بوزيتي، كالينا بونتشيفا

يتناول هذا القسم الاستجابات الانتخابية المُصمَّمة لحماية الناخبين ونزاهة ومصداقية الانتخابات، من خلال تدابير تهدف إلى كشف وتعقب ومُكافحة التّضليل الذي ينتشر خلال الحملات الانتخابية. ويُهدّد هذا التّضليل العملية الديمقراطية بشكل عام في عدد متزايد من البلدان في العالم (UNESCO, 2019).

ويلقي هذا القسم الضوء على المُبادرات التي أطلقتها، إما وسائل الإعلام أو المنظمات غير الحكومية، وأحياناً الهيئات الانتخابية نفسها. وهدفها هو منع تهديد الانتخابات وتقويض الديمقراطية، والحفاظ على المعايير العالمية لنزاهة الانتخابات التي تنطبق على جميع البلدان في كافة مراحل الدورة الانتخابية –من الفترة التي تسبق الانتخابات، مروراً بالحملات الانتخابية، ثم عملية الاقتراع، وفي أعقابها (Norris et al., 2019).

وقد تم البحث في الاستجابات السياسية والقانونية التي تتخذها الدولة في القسم 5.1، فيما يتناول القسم 5.2 الحملات الوطنية والدولية لمُكافحة التّضليل. ولا شك في أنّ وقائع عصر الإنترنت تُعقّد المعايير التي كانت متبعةً قبل هذا العصر على غرار المعايير الواردة في «دليل جوانب الانتخابات القانونية والتقنية وتلك المتعلقة بحقوق الإنسان» (United Nations, 1994).

> ”لا بدّ أن يتمتع استخدام وسائل الإعلام لأغراض الحملات بالمسؤولية من حيث المُحتوى، حتى لا يُسمح لأيّ حزب بإطلاق البيانات الكاذبة أو التشهيرية أو العنصرية، أو التي تُحرّض على العنف. كذلك لا يجب أن يُسمح بقطع الوعود غير الواقعية أو المخادعة، ولا أن تقوم الأحزاب بتعزيز التوقعات الزائفة عن طريق استخدام وسائل الإعلام.“

1.3.5 ما هي الأمور ومن هي الجهات التي تستهدفها الاستجابات للتضليل الانتخابي؟

لقد أدى التشكيك في مصداقية بعض وسائل الإعلام الإخبارية وانتشار الأدوات الرقمية التي تسهّل إنتاج مواد إعلامية ملفقة، تقلّد الصحافة الموثوقة؛ إلى زيادة انتشار التّضليل خلال الفترات الانتخابية [150] (Ireton & Posetti 2018; Norris et al., 2019). وصحيح أن الناس يميلون إلى تصديق بعض الأكاذيب والخرافات التي تنتشر عبر حملات التّضليل المنظمة ولكنّ الضرر الرئيسي يكمن في أنّ انتشار هذه المعلومات يؤدي إلى تآكل قدرة المواطنين على تمييز الحقيقة، مما يؤدي بدوره إلى تحويل الانتخابات إلى مسابقات شعبية لا تحتاج إلى معلومات صحيحة ومدققة. هذا من شأنه أن يسهم في تآكُل مفهوم أساسي في الحياة الديمقراطية وهو قيام الناخبين باتخاذ خيارات سياسية عقلانية مبنية على المعرفة.

[150] يُرجَى مراجعة النقاش حول الثقة في الفصل 7.1.

أما المعلومات المُضلّلة التي تنتحل صفة المُحتوى الإخباري الشرعي فغالباً ما يتم كشف زيفها بشكل سريع. ومع ذلك، فإنّ هدفها لا يقضي بالضرورة باختلاق معتقدات مبنية على الزيف، بل «النَّيل من المعتقدات والقناعات القائمة... بهدف زعزعة الاستقرار، والتشكيك بالقوى القائمة والقوى المضادة على حد سواء؛ لجعلنا نفقد الثقة في مصادرنا ولزرع البلبلة» (Eco, 2014). وفي حين ينطبق ذلك على التّضليل في ما يخص مجموعة من القضايا (كالتلقيح، وتغيُّر المناخ، والهجرة)، فقد يكون له ايضاً أهمية مباشرة خلال الانتخابات.

فعلى سبيل المثال، في سياق الاستفتاء على عضوية بريطانيا في الاتحاد الأوروبي في عام 2016 المعروف باسم تصويت «بريكست– Brexit»، يؤكد بعض الباحثين أن تعرض الناخبين للمعلومات المُضلّلة على وسائل التواصل الاجتماعي لعب دوراً كبيراً في تحديد نتائج الاستفتاء (Parkinson 2016; Read, 2016; Dewey 2016). في الوقت نفسه أشار باحثون آخرون إلى أن تعقيد الوضع السياسي والاستقطاب الشديد لعبا دوراً أهم (e.g. Benkler et al., 2018; Allcott & Gentzkow, 2017; Guess et al., 2018b)، في حين سلّط البعض الضوء على دور التغطية الصحفية المنحازة (Davis, 2019; Freedman, 2016). وأشار أحد تقييمات السياسة الخارجية إلى فشل المُساءلة الصحفية في تطبيق المعايير المهنية في ما يتعلق بقول الحقيقة (Barnett, 2016):

> ”فشلت وسائل الإعلام التقليدية فشلاً ذريعاً (...) فقد انغمست معظم الصحافة الوطنية في المملكة المتحدة في تشويه الحقيقة ونشر أنصاف الحقائق والأكاذيب، أي في حملة دعائية شرسة تمّت فيها التضحية بالحقائق والتحليل الواعي لصالح الأهداف المؤدلجة للمحررين وأصحاب هذه الوسائل... وكذلك فإنّ تشكيكهم المُعادي لأوروبا كان له دور في وضع أجندات المذيعين.“

ومن العوامل الأخرى فشل القواعد الموضوعية داخل المؤسسة الصحفية بسبب المفهوم الخاطئ الذي مفاده أن لابد من إعطاء كلا الجانبين من النقاش الوزن نفسه (أي من يدعم خروج المملكة المتحدة من الاتحاد الأوروبي، ومن يدعم بقاءها فيه)، بدلاً من تقييم الجانبين على أساس أدلة تثبت عمل كل منهما للمصلحة العامة. ومع ذلك، ينبغي الاعتراف بأن وسائل الإعلام كانت تعتمد استراتيجية تواصل سياسية «مواليةً لخروج بريطانيا» ومصممةً «على زعزعة الخطاب والسيطرة على رسائل وسائل الإعلام عبر الاجتذاب العاطفي للناخبين». وعندما اختلطت هذه الاستراتيجية بالمعلومات المُضلّلة كان لها تأثير قوي على العملية الديمقراطية (Beckett, 2016).

ومثال آخر يؤكد الحاجة إلى مُكافحة التّضليل ذو الصلة بالانتخابات على فيس بوك وغيره من مواقع التواصل هو الانتخابات الرئاسية الأمريكية في عام 2016. ففيما أكد العلماء على أنّ الاستقطاب كان موجوداً قبل ذلك في السياسة الأميركية، فقد تمّ الاعتراف أيضاً بالدور المهم الذي تضطلع به حملات التّضليل المنظمة (مثل: فضيحة كامبريدج أناليتيكا) في المعادلة الأوسع (Benkler et al., 2018; Allcott & Gentzkow, 2017; Guess et al., 2018).

ومن المخاوف الأخرى التي لابد من معالجتها؛ المعلومات المُضلّلة المرتبطة بالدعاية السياسية، وقدرتها على التأثير على الناخبين بطريقة غير شريفة. ويمكن توزيع هذا المُحتوى كرسائل على شبكات التواصل الاجتماعي، وعلى تطبيقات الدردشة المغلقة، وعلى شكل «ميمات Memes»، ومقاطع

فيديو، وصور تهدف إلى اقناع الناخبين وتعبئتهم، أو إلى قمع الناخبين وثيهم عن الاقتراع (Wood & Ravel, 2018). ويتم تصميم هذه الإعلانات لتؤثر على آراء الناس السياسية وعلى إقبال الناخبين أو ثنيهم عن الاقتراع. ويدفع المعلن لإحداث هذه الآثار، ويمكنه نشر هذه الاعلانات من خلال الاستهداف الدقيق على وسائل الإعلام الاجتماعية ومُحركات البحث. وقد تبدو بعض الإعلانات السياسية وكأنها محتويات عضوية أو إعلانات أصلية؛ يصعب تتبعها أكثر ولا يمكن مقابلتها بسهولة بسرديات مضادة. تجدر الإشارة أيضاً إلى أنّ الشروط المفروضة في العديد من البلدان على وضع إعلانات سياسية على مواقع التواصل الاجتماعي هي أيسر عموماً من الشروط المفروضة على إعطاء ترخيص بالبث.

وكان الإنفاق الإعلاني السياسي يتصاعد بشكل كبير قبل الانتخابات الرئاسية الأميركية في عام 2020 حيث تنبأت شركة تسويق رقمي أنّ إجمالي الإنفاق على إعلانات الحملات الانتخابية سوف يقفز بنسبة 63% مقارنةً بانتخابات عام 2016 ليصل إلى 6.89 مليار دولار أميركي (eMarketer, 2020). ووفقاً لهذا التقرير، دفعت البيئة السياسية المتحزبة إلى حد كبير بعدد أكبر من الأمريكيين إلى التبرُّع لمرشحيهم مقارنةً بالانتخابات السابقة، ممّا أدى بدوره إلى تحويل المزيد من الأموال نحو الإعلانات السياسية. وفي حين كان من المتوقع أن يستحوذ التلفزيون على الحصة الأكبر من الإعلانات السياسية (أي 66% من المجموع)، كان من المتوقع أن ينمو الإعلان الرقمي خاصةً على فيس بوك الذي يعتبر المنصة الرئيسية بأكثر من 200% مقارنةً بالانتخابات الرئاسية السابقة، وفقاً للمصدر نفسه. ويجتذب فيس بوك المعلنين السياسيين لتوفيره نطاق وصول واسع ولقدراته في استهداف الناخبين المُتبارَى في سبيل نيل تصويتاتهم (Harding-McGill & Daly, 2020)، ولسهولة استخدامه كذلك.

وقد أدت هذه العوامل مجتمعةً إلى ظهور «حاضنات مأجورة للحسابات الدُّمى» يديرها وكلاء التّضليل من وحدات الدعاية التابعة للدول والجهات التي تسعى إلى تحقيق الاستفادة وشركات العلاقات العامة التي بدأت تتخصص في إنشاء شبكات تضليل منظمة من خلال استخدام مجموعة من الأساليب التي تتخطى الإعلان. وهي معروفة باسم «شركات العلاقات العامة السوداء» (Silverman et al., 2020; Bell & Howard, 2020). وهناك مخاوف متزايدة من الدور الذي قد يلعبه مروجو التّضليل في السياقات الانتخابية. ويكمن الخطر في أن هذه الشبكات، والتي تتخصص أيضاً في «الدعاية الشعبية الزائفة»، مصممة لتقليد المواطنين والحركات السياسية العضوية، وبالتالي لإضفاء قشرة زائفة من الشرعية تسهّل انتشار محتواها بشكل أوسع من الإعلانات السياسية المعروفة.

وقد ظهر مثال آخر على الهويات والسلوكيات المخادعة على الإنترنت خلال الانتخابات العامة في المملكة المتحدة في عام 2019 عندما تم تغيير اسم حساب مقر الحملة الانتخابية للحزب المحافظ على تويتر من CCHQPress إلى FactCheckUK. كما وتم تغيير الصورة الرمزية للحساب لتبدو وكأنها الصورة الرمزية لإحدى المنظمات التي تعمل في مجال التحقُّق من الوقائع خلال مناظرة تلفزيونية بين رئيسي الحزبين». وكل تغريدة نشرت خلال هذه المناظرة بدأت بكلمة «FACT» أي حقيقة. وبعد النقاش، أُعيد تغيير اسم الحساب والصورة الرمزية. وقد دافع حزب المحافظين الذي انتصر في نهاية المطاف عن تصرفه هذا في حين اتهمه تويتر بتضليل الرأي العام، وهو رأي رددته منظمة FullFact وهي منظمة مستقلة للتحقق من الوقائع (Perraudin, 2019). وقد أكد استخدام التحقُّق من الوقائع لتحقيق مكاسب سياسية خلال الحملات الانتخابية على قيمة هذه الخدمات كأدوات ثقة، في حين أثار أيضاً مخاوف كبيرة داخل مجتمعات التحقُّق من الوقائع ومجتمع الإعلام.

وقد استجابت الجهات الصحفية لهذه الأشكال من التّضليل المتصل بالانتخابات عبر تحقيقات صحافية وتحليلات جنائية للبيانات (Ressa 2016; Silverman et al., 2020).[151] وقد بنت منظمات التحقُّق من الوقائع على هذه التقاليد في مشروعات خاصة بالانتخابات (انظر أدناه).

وتضمنت الاستجابات المتخذة على مستوى الدولة دعوات لتشديد تنظيم الإعلان السياسي، وشبكات الدعاية، واستهداف الناخبين في بعض السياقات (Dobber et al., 2019; Kelly, 2020b)، ولكنها دعت إلى اعتماد أنظمة أقل تشدداً في سياقات أخرى (@TeamTrump 2019). وقد يُفسّر هذا الاختلاف في المقاربات جزئياً باحتمال الخسارة أو الربح بالنسبة إلى الأحزاب السياسية الحاكمة.

وإلى جانب الصحفيين، فإنّ الجهات الرئيسية الأخرى المشاركة في الاستجابة للتضليل الانتخابي هي شركات التواصل عبر الإنترنت. فخلال عام 2020، اندلع خلاف علني بين تويتر وفيس بوك على خلفية تباين في مقاربات التحقُّق من الوقائع وتحديد التّضليل المرتبطة بادعاءات الرئيس الأمريكي حينذاك حول العملية الانتخابية (Smith, 2020b). ويقع التحقُّق من الوقائع والتحذير من ادعاءات الرئيس السّابق دونالد ترامب ضمن المبادئ التوجيهية المتعلقة بمسألة رصد التّضليل الانتخابي والتحقُّق منه في الشركتين. وفي أيار/مايو 2020، اتخذ تويتر خطوةً غير مسبوقة بوضع إشارة تحذير على التغريدات ذات الصلة (NPR, 2020). وقد اعترض على هذه الخطوة الرئيس التنفيذي لفيس بوك مارك زوكربيرغ بشدة، بحجة أن الشركات الخاصة لا ينبغي أن تلعب دور «الحكم في الحقيقة» (Halon, 2020). لكن، وكما جاء في الفصل 7.1، إنّ تجنب لعب دور الحكم في الحقيقة لا يعني عدم اتخاذ إجراءات لمُكافحة الترويج للأكاذيب الواضحة (Kaye, 2020b). ورداً على قرار تويتر في ما يتعلق بتطبيق سياسته حول الاستجابة للتضليل الانتخابي، أعلن الرئيس الأمريكي السّابق على الفور (وعبر تويتر) أنه سيعمد إلى «التشدد في تنظيم» شركات التواصل الاجتماعي أو حتى «إغلاقها» بموجب أمر تنفيذي (Smith & Shabad, 2020). وقد أدانت منظمات المجتمع المدني التي تعمل في مجال حُريّة التّعبير هذا التهديد، فيما اعتبر آخرون أنّه لا يمكن تطبيق الأمر التنفيذي من دون تعديل قانوني (Article 19, 2020b).

2.3.5 مَن هي الجهات التي تسعى الاستجابات الانتخابية للتضليل إلى مُساعدتها؟

تهدف الاستجابات الانتخابية إلى حماية الناخبين من التعرض للتضليل والتقليل من احتمال تأثير التّضليل على وجهات نظرهم السياسية وعلى تصويتهم بطرق لم تكن لتحدث لولا تعرضهم لهذا التّضليل.

في سياق الانتخابات، تم استخدام الإعلان السياسي (بما في ذلك الرسائل الشخصية والمستهدفة بشكل فردي) على نطاق واسع من قبل الأحزاب السياسية والمرشحين بهدف التأثير على الناخبين. فعلى سبيل المثال، خلال الاستفتاء على عضوية بريطانيا في الاتحاد الأوروبي في عام 2016، استخدمت الحملة الداعمة لخروج بريطانيا #VoteLeave إعلانات تعتمد الاستهداف؛ تتضمن معلومات مضللة حول التكلفة الأسبوعية لعضوية بريطانيا في الاتحاد الأوروبي وحول التحضير لانضمام تركيا للاتحاد الأوروبي (Cadwalladr, 2018). وفي عدد من الحالات، أُعطيت المعلومات المُضلّلة شكل إحصاءات، مما

[151] https://www.theguardian.com/news/series/cambridge-analytica-files

أثار مسائل معقدة تتعلق باحتساب التكاليف والفوائد. وعلماً أنّ المنافسات السياسية تنطوي دائماً على استخدام انتقائي للاحصاءات، تظهر مناطق رمادية في عملية تحوّل الحملات الانتخابية إلى أعمال تضليل، على الرغم من أنّ الإحصاءات الملفقة ليست حتماً من الأعمال التثقيفية التي تسعى إلى تقديم المعلومات الصحيحة. ويمكن أيضاً استخدام عبارة التّضليل/المعلومات المُضلّلة (وعبارتي المعلومات المغلوطة و«الأخبار الزائفة») كسلاح لتصنيف بعض الادعاءات الحقيقية على أنها غير صادقة وغير نزيهة. وتؤكد هذه التحديات على أهمية وتعقيد استجابات التصدي للتضليل الانتخابي.

ومن الطرق الأخرى التي تساعد فيها الاستجابات الانتخابية الناخبين على التعامل مع التّضليل هو كشف الجهات الكامنة وراء هذه المشكلة. فعلى سبيل المثال، لا يدرك العديد من الناخبين دائماً أن الوسيلة الرئيسية لاستهدافهم بالمعلومات المُضلّلة هي من خلال الحسابات الآلية (اي روبوتات الإنترنت والسايبورج «شخصية خيالية نصف بشريّة-نصف آليّة). ومع ذلك هناك حالات مدروسة جيداً تؤكد حصول هذا الأمر خلال الانتخابات الرئاسية الأمريكية في عام 2016، والانتخابات الرئاسية الفلبينية في عام 2016، والاستفتاء على عضوية بريطانيا في الاتحاد الأوروبي والانتخابات الرئاسية الفرنسية في عام 2017. وقد ثبت بشكل خاص أنّ الروبوتات السياسية تحاول التأثير على رأي الناخبين، عبر مهاجمة القادة السياسيين أو الصحفيين مثلاً، على الرغم من أن بعض الأدلة يشير إلى أن الروبوتات لا تغيّر رأي الناخبين في بعض الأحيان (Howard et al., 2018b). ولكن الحقيقة هي أنه خلال الانتخابات تم استخدام عدد كبير من الروبوتات والحسابات الدمى المُختلقة لنشر التّضليل والشائعات السياسية (Ressa 2016; Phillips & Ball, 2017; Howard et al., 2018; Gorrell et al., 2018; Howard & Kollanyi, 2016). ويُعد كشف هذه الظواهر جزءً من الاستجابات الانتخابية التي يمكن أن تعمل على توعية الناخبين بعمليات التّضليل السرية و«الحيل الرخيصة» التي تستهدف تقويض معايير وقواعد الحملات الانتخابية النزيهة.

3.3.5 ما هي المخرجات التي تنشرها الاستجابات الانتخابية؟

قد تتضمن مخرجات الاستجابات الانتخابية للتضليل عمليات كشف في الوقت الحقيقي، وعمليات كشف الزيف ونشر المُحتوى المضاد والتقييمات وردود الأفعال. كما وقد تتضمن حملات تسعى إلى تثقيف الناخبين، ولوائح تنظم السلوك الانتخابي.

ويتم التطرق للاستجابات الانتخابية كفئة مستقلة في هذا التقرير نظراً لأثر التّضليل الكبير على العمليات الديمقراطية وحقوق المواطنين خلال الانتخابات. ومع ذلك، ونظراً لطبيعة هذه الفئة من الاستجابات، فإنها غالباً ما تتضمّن مزيجاً من استجابات الرصد والتحقُّق من الوقائع، والاستجابات الخاصة بتنظيم المُحتوى والاستجابات التقنية والتثقيفية وغيرها من الاستجابات، والتي تم تصنيفها بشكل منفصل في التصنيف الوارد في هذا التقرير. وتتم الإشارة إليها في هذا القسم عند الاقتضاء. ولذلك، فإن مخرجات الاستجابات الانتخابية هي أساساً مجموعة فرعية من مخرجات الفئات الأخرى من الاستجابات (على سبيل المثال: التحقُّق من الوقائع الانتخابية ومحفوظات الإعلانات الانتخابية).

4.3.5 مَن هي الجهات الفاعلة الرئيسية في مجال الاستجابات الانتخابية ومن يمولها؟

أ. التحقُّق من الوقائع السياسية في الولايات المتحدة:

يرتبط تاريخ التحقُّق من الوقائع السياسية في الولايات المتحدة بقوة، بتغطية الانتخابات الرئاسية وكمية الأكاذيب التي تُنشر خلال الأحداث العاجلة. حتى الآن، يُركّز معظم التحليل المفصل لممارسات التحقُّق من الوقائع السياسية على الولايات المتحدة التي كانت مصدر نشوء هذه الممارسة (Birks, 2019). والحقيقة أن حركة التحقُّق من الوقائع في الولايات المتحدة ظهرت رداً على ما اعتُبر فشل لوسائل الإعلام في الكشف عن الأكاذيب خلال الحملات الانتخابية (Spivak, 2010).

وكانت منظمة التحقُّق من الوقائع الأولى المستقلة هي Spinsanity[152]، التي تأسست في عام 2001 (Graves, 2013) ونشطت خلال الحملة الانتخابية الرئاسية لعام 2004، فنشرت أكثر من 400 مقالة. بعد ذلك، قبيل الانتخابات الرئاسية الأمريكية لعام 2004، تم إطلاق FactCheck.org التي تم تعريفها على أنها «منظمة غير متحزبة وغير ربحية تدافع عن المستخدمين والناخبين» وتهدف إلى مراقبة كل من الأحزاب السياسية الرئيسية، والبرامج الحوارية، والإعلانات التلفزيونية والمواقع الإلكترونية الرسمية والنشرات الصحفية ومحاضر المؤتمرات الصحفية.[153] ومن المبادرات البارزة الأخرى مبادرة The Fact Checker، التي أطلقتها صحيفة واشنطن بوست قبل انتخابات عام 2008 (Kessler, 2017). وقد أطلقت نظاماً جديداً يصنّف المعلومات المُضلّلة باستخدام سلّم اسمه «بينوكيو» وفيه 4 درجات من 1 إلى 4.

ويُعتبر إنشاء Politifact لانتخابات عام 2008 تطوراً رئيسياً آخر في مجال التحقُّق من الوقائع السياسية، وهي أكبر وسيلة مستقلة للتحقق من الوقائع في الولايات المتحدة (Aspray & Cortada, 2019; Drobnic Holan, 2018). والجدير بالذكر أن هذه المبادرة أصبحت بارزةً لنوعية تحققها من الوقائع، وخاصةً نظام تصنيف الحقيقة فيها المسمى Truth-O-Meter (مقياس يتراوح بين صحيح، في الغالب صحيح، نصف صحيح، خاطئ في الغالب، خاطئ، كذب محض[154]). وقد أصبح سُلّم التصنيف Truth-O-Meter معلماً مميزاً من Politifact التي فازت بجائزة «بوليتزر» في عام 2009 لتغطيتها الحملة الرئاسية لعام 2008 (Adair, 2018). وكان موقع FactCheck.org مرشحاً أيضاً للفوز بهذه الجائزة (Graves, 2013).

وفي عام 2010، وسّعت Politifact نشاطها من خلال منح ترخيص علامتها التجارية ومنهجيتها لشركاء من وسائل الإعلام في الولايات المتحدة الأمريكية. وبعد ثلاث سنوات، أطلقت Politifact مبادرة Punditfact للتحقق من دقة أقوال الخبراء وكُتّاب الأعمدة والمدونين والمحللين السياسيين والمضيفين والضيوف من البرامج الحوارية، وأعضاء آخرين من وسائل الإعلام (Hollyfield, 2013).

في الوقت الحاضر، تُعتبر FactCheck.org وPolitifact التابعة لصحيفة واشنطن بوست وSnopes الوسائل الأهم للتحقق من الوقائع في الولايات المتحدة (Graves 2013; Aspray & Cortada, 2019). وهي تواجه

152 http://www.spinsanity.org/about/

153 https://www.factcheck.org/spindetectors/about/

154 https://www.politifact.com/truth-o-meter/

تحديات هائلة، حيث أن هناك، كما يؤكد الباحثون، نوعين من البيئات المحيطة بوسائل الإعلام في الولايات المتحدة. الأولى هي «البيئة الإعلامية الانعزالية اليمينية»، والتي تظهر عليها «كل مظاهر غرف الصدى التي تسعى إلى زرع التشدد في صفوف السكان وتقويض قدرتهم على التمييز بين الحقيقة والخيال، وزعزعة ثقتهم في المؤسسات». والبيئة الأخرى، والتي تمثل غالبية وسائل الإعلام، هي «أقرب إلى نموذج الحيّز العام المتصل ببعضه البعض» (Benkler et al., 2018). وفي هذا النظام البيئي المزدوج، ينظر «اليمين الانعزالي» إلى مواقع التحقُّق من الوقائع على أنها منحازة بشكل منهجي فلا يثق بها أو يصدقها (المرجع نفسه).

ب. التحقُّق من الوقائع السياسية في أوروبا:

بدأ التحقُّق من الوقائع في أوروبا كرد فعل على التّضليل السياسي عندما قامت قناة الأخبار البريطانية 4 News بإنشاء موقع للتدوين في عام 2005 لتغطية الانتخابات البرلمانية (Graves & Cherubini, 2016). تلا ذلك إنشاء مواقع فرنسية مماثلة على غِرار موقع Désintox التابع لصحيفة Libération في عام 2008، وموقع Les Décodeurs التابع لصحيفة Le Monde في عام 2009. وقد استوحى الموقعان من موقعي Politifact وFactCheck.org وهما يهدفان إلى التحقُّق من الوقائع المتعلقة بالشخصيات السياسية والعامة وبالحملات الانتخابية. وقد أُطلقت المنظمة الخيرية البريطانية FullFact.org في عام 2009، بقصد «مُكافحة المعلومات السيئة» وانضم إليها أيضاً موقع Reality Check التابع لقناة BBC (Birks, 2019). وفي هولندا، انطلق مشروع التحقُّق من الوقائع Nieuwscheckers في العام نفسه، وذلك في مدرسة الصحافة والإعلام الجديد في جامعة ليدن. وتوسع التحقُّق من الوقائع بسرعة في أوروبا، خاصةً في ما يخص الانتخابات. فمن أصل 34 منظمةً عاملةً في مجال التحقُّق من الوقائع في 20 دولةً أوروبيةً في العام 2016، ارتفع هذا الرقم في بداية عام 2020 إلى نحو 66 منظمةً ناشطةً في 33 بلداً في المنطقة، وفقاً لبيانات مختبر المراسلين في جامعة ديوك.

ولطالما كانت الانتخابات الرئاسية أو العامة حافزاً لتوسيع نطاق حركة التحقُّق من الوقائع إما عن طريق إنشاء تجارب «حياة حقيقية»، أو عمليات أكثر استدامة. فعلى سبيل المثال، بدأت الصحافة الفرنسية بالتحقُّق من الوقائع خلال حملة الانتخابات الرئاسية في عام 2012 (Bigot, 2019).

أما في النمسا وإسبانيا وإيطاليا، فقد دخل التحقُّق من الوقائع في الإعلام التقليدي عبر قنوات التلفزيون. فقد أطلقت المؤسسة النمساوية للإرسال ORF في عام 2013 برامج تسعى إلى التحقُّق من الوقائع المتصلة بالسياسيين مباشرةً على الهواء، في الفترة التي تسبق الانتخابات العامة. وفي العام نفسه، في إسبانيا، بثت شركة Objetivo برنامجاً في وقت الذروة على تلفزيون La Sexta هدفه التحقُّق من الوقائع الخاصة بالسياسيين في ظل الأزمة المالية الإسبانية. وبعد بضعة سنوات، تم بث برنامج مُماثل أعدته Pagella Politica في إيطاليا على القناة الوطنية، TV RAI2.

وفي عام 2018، اقترح تقرير أعده فريق الخبراء رفيع المستوى حول التّضليل التابع للمفوضية الأوروبية؛ عدة استراتيجيات لمُكافحة التّضليل وحماية الانتخابات في الاتحاد الأوروبي وفي الدول الأعضاء، عبر تعزيز الشفافية في الدعاية السياسية على سبيل المثال وتطوير أدوات تسمح بتمكين المستخدمين وتعزيز الدراية الإعلامية. في وقت لاحق من ذلك العام، أعلنت المفوضية الأوروبية عن تدابير لضمان إجراء انتخابات حرة ونزيهة للبرلمان الأوروبي في أيار/مايو 2019 (European Commission, 2018b).

وتشـمل تلـك التدابيـر توصيـات للـدول الأعضـاء بإنشـاء شـبكة وطنيـة مـن السُّـلطات المختصـة لكشـف المخاطـر المتعلقـة بالأمـن السـيبراني والتّضليل والاسـتجابة لهـا، ولزيـادة شـفافية الإعلانـات والاسـتهداف علـى الإنترنـت، وتشـديد الضوابـط علـى تمويـل الأحـزاب السياسـية الأوروبيـة.

ج. التحقُّق من الوقائع في بقية العالم:

فـي منطقـة آسـيا والمحيـط الهـادئ، سـجلت قاعـدة بيانـات مختبـر المراسـلين فـي جامعـة ديوك 47 منظمـةً تعمـل فـي مجـال التحقُّـق مـن الوقائـع. علـى النحـو المبيـن أدنـاه، انتشـر التّضليـل الإعلامـي فـي الهنـد وإندونيسـيا والفلبيـن وجمهوريـة كوريـا فـي مـا يتعلـق بالانتخابـات بشـكل كبيـر مقارنـةً باليابـان وسـنغافورة وأسـتراليا. وفـي أجـزاء أخـرى مـن العالـم، تطـوّرت المبـادرات الحديثـة للتحقـق مـن الوقائـع فأصبحـت مبـادرات لكشـف زيـف المعلومـات المُضلّلـة أكثـر منهـا لكشـف الادعـاءات والخطابـات السياسـية، لذلـك يتـم تناولهـا فـي القسـم 4.1.

وقـد بـدأت جميـع الأحـزاب السياسـية فـي الهنـد باسـتخدام وسـائل التواصـل الاجتماعـي خـلال الحمـلات الانتخابيـة لعـام 2014، مـع التركيـز علـى اسـتهداف الأشـخاص الذيـن يقترعـون للمـرة الأولـى (Kaur & Nair, 2018). وفـي الآونـة الأخيـرة، تطـوّر تطبيـق واتسـاب ليصبـح قنـاة التّضليـل الرئيسـية (Kajimoto & Stanley, 2019) فـي الهنـد.

وجـزء كبيـر مـن المعلومـات المُضلّلـة التـي تـم كشـفها فـي الهنـد هـي سياسـية وتتعلـق إمـا بخلافـات محليـة أو بالتوتـر مـع الجـارة باكسـتان. ومـن الجديـر بالذكـر أنـه فـي حملـة انتخابـات الجمعيـة التشـريعية فـي نيودلهـي فـي شـباط/فبراير مـن عـام 2020، نشـر أعضـاء حـزب سياسـي معيّـن مقطعـي فيديـو مزيفيـن علـى تطبيـق واتسـاب مـع رسـائل تسـتهدف منافسـهم السياسـي (Christopher, 2020).

وفـي إندونيسـيا، ينتشـر التّضليـل فـي كثيـر مـن الأحيـان خـلال الانتخابـات المهمـة مسـتغلاً الاختلافـات الدينيـة والعِرقيـة (Kajimoto & Stanley, 2019). ويُسـمى الفاعلـون الرئيسـيون «الأبـواق السياسـية» وهـم يسـعون إلـى تدعيـم موقفهـم الانتخابـي وتقويـض حمـلات منافسـيهم عبـر خطـاب الكراهيـة والخطـاب شـديد التحـزب أو عبـر التّضليـل الدينـي والقائـم علـى أسـاس العِـرق.

والفلبيـن هـي الأُخـرى دولـة تعانـي مـن انتشـار التّضليـل فـي الخطـاب السياسـي عبـر الإنترنـت، لا سـيما منـذ حملـة الانتخابـات الرئاسـية فـي عـام 2016. وكثيـراً مـا يلقـى باللـوم علـى مـا يُعـرف «بالمسـتفزين ذوو المشـاعر الوطنيـة» وشـبكات الإنترنـت المنظمـة «لانعـدام الأخـلاق» والخـروج علـى حـدود اللياقـة اللذيـن يشـوبان الخطـاب السياسـي منـذ تلـك الفتـرة (Ressa 2016, Ong & Cabañes 2018).

ويُقـال إنّ بعـض جهـود التحقُّـق مـن الوقائـع التـي تضطلـع بهـا الـوكالات الإخباريـة والمنظمـات غيـر الحكوميـة فـي الفلبيـن تفشـل فـي معالجـة الأسـباب الكامنـة وراء التّضليـل لأنهـا لا تعالـج الهياكـل «الاحترافيـة والمؤسسـية والحوافـز الماليـة التـي تعمـل علـى تطبيـع ومكافـأة عمـل المسـتفزين علـى الإنترنـت» (Ong & Cabañes, 2018). وقـد تـم التبليـغ بانتظـام عـن وجـود «حاضنـات النقـر المأجـورة» وعـن ممارسـة الدعايـة الشـعبية الزائفـة، خصوصـا علـى فيـس بـوك منـذ عـام 2016 فـي الفلبيـن.

أمّا في جمهورية كوريا، فقد قامت كافة الصحف والإذاعات تقريباً بإطلاق مبادرات للتحقق من الوقائع خلال الانتخابات الرئاسية في عام 2017 (Kajimoto & Stanley, 2019). وكان هدفها التصدي لانتشار المعلومات المُضلّلة، بما في ذلك عبر مسعى تعاوني مع المؤسسات الأكاديمية. وأطلقت جامعة سيول الوطنية[155] منصة تحقق من الوقائع اسمها SNU Factcheck تخوّل 26 وسيلةً إخبارية؛ التحقُّق بين بعضها البعض من المعلومات المختلف عليها. ومن الأمثلة الأخرى في المنطقة مركز FactCheck[156] وMyGoPen اللذين يتحققان من المعلومات المُضلّلة على تطبيق المراسلة LINE.

د. استجابات وسائل الإعلام التعاونية في مجال الانتخابات:

نظراً للحجم الهائل من المعلومات المُضلّلة وبيانات المرشحين على الإنترنت التي تحتاج إلى التحقُّق خلال الانتخابات، بدأت مؤسسات إعلامية عديدة في حشد مواردها في مبادرات تعاونية مُعد لها بشكل جيد، بعضها ذات طبيعة وطنية بينما بعضها الآخر ذات طبيعة دولية. ويناقش هذا القسم بعض الأمثلة البارزة.

الاستجابات التعاونية القُطريّة

شكلت مبادرة Electionland المسعى المشترك الأول في الولايات المتحدة في عام 2016، وقد أطلقتها Propublica بالتعاون مع مختبر Google للأخبار، وWNYC، وFirst Draft، وشبكة U.S. Today Network، وNews Univision، ومدرسة الصحافة في جامعة مدينة نيويورك. والهدف من هذه المبادرة رصد المعلومات المُضلّلة المنشورة على وسائل التواصل الاجتماعي حول انتخابات عام 2016. حيث تضمّن المشروع 600 طالب صحافة وأكثر من 400 صحفي في جميع أنحاء الولايات المتحدة (Bilton, 2016; Wardle, 2017b).

وفي **أوروبا**، كانت مبادرة CrossCheck France (بتمويل من مختبر أخبار Google) من أوائل مشاريع الصحافة التعاونية التي عملت على كشف الروايات والتعليقات والصور وأشرطة الفيديو غير الصحيحة حول المرشحين والأحزاب السياسية وكافة القضايا الأخرى ذات الصلة بالانتخابات التي انتشرت على الإنترنت خلال حملة الانتخابات الرئاسية الفرنسية عام 2017. وقد تضمّنت هذه المبادرة أكثر من 100 صحفي من 30 وسيلة إعلام فرنسية ودولية إلى جانب بعض الأكاديميين وشركات التكنولوجيا. في المجموع، تم إنتاج 67 منشوراً لكشف الزيف وذلك على موقع CrossCheck، وكذلك على مواقع الشركاء في غرفة الأخبار (Smyrnaios et al., 2017). واجتذبت هذه التجربة التعاونية الرائدة في كشف الزيف 336000 زائر (95% منهم فرنسيون) (المرجع نفسه).

وقبل الانتخابات العامة في المملكة المتحدة عام 2017، أنشأت المنظمة غير الربحية First Draft مشروع UK CrossCheck[157]، الذي خصص مساحةً للصحفيين البريطانيين للعمل على التحذير من الزيف، وعلى تيسير التعاون في مسائل الإبلاغ بهدف الكشف المُبكّر، وعلى التحقيق في المُحتوى والسلوكيات المشتبه بها على الإنترنت. وفي تصنيف الاستجابات الواردة في هذا التقرير، يُعتبر مشروع CrossCheck UK استجابةً استقصائيةً بما أنّه يركز على السرديات والسياقات المُضلّلة لا على تقييم الادعاءات الفردية وتقرير صحتها من عدمها. ومصادر تمويل هذه النسخة من CrossCheck غير واضحة.

155 http://factcheck.snu.ac.kr/

156 https://tfc-taiwan.org.tw/

157 https://firstdraftnews.org/project/crosscheck-uk/

وفـي ذلـك الوقـت، ظهـرت الحاجـة إلـى وضـع منهجيـة ومعرفـة وأدوات مشـتركة بيـن الاسـتجابات للتضليـل، التـي تركـز علـى وسـائل الإعـلام.[158] ولذلـك كانـت مسـاهمة First Draft ذات أهميـة كبـرى؛ كمـا وتـم تطويـر أدوات مبتكـرة وُضعـت خصيصـاً لدعـم العمليـات التعاونيـة فـي مجـال التدقيـق والتحقُّـق مـن الوقائـع (Mezaris et al., 2019).

وقـد قامـت مبـادرة إسـبانية تُعـرف باسـم Comprobado باعتمـاد منهجيـة CrossCheck التعاونيـة بالتعـاون مـع Maldita.es لمُكافحـة التّضليـل خـلال الانتخابـات العامـة فـي البـلاد فـي عـام 2019. بالإضافـة إلـى التحقُّـق مـن الوقائـع والتحقيـق فـي التّضليـل السياسـي؛ حيـث تـم التصـدي لتحـدٍّ جديـد هـو التّضليـل علـى تطبيقـات المراسـلة (واتسـاب علـى وجـه الخصـوص). وقـد اتبعـت مبـادرة Comprobado ضوابـط جـودة صارمـة فـي مـا يخـص التدقيـق عـن طريـق اشـتراط موافقـة ثلاثـة علـى الأقـل مـن أعضـاء المشـروع البالـغ عددهـم 16. وبنـاءً علـى الـدروس المسـتفادة مـن المبـادرات السـابقة، وبعنايـة فائقـة، قامـت Comprobado باختيـار المُحتـوى واسـع الانتشـار الـذي لابـد مـن كشـفه، والطريقـة للقيـام بذلـك لتجنـب انتشـار المعلومـات المُضلّلـة.

وفـي عـام 2018، بـدأت المبـادرات التعاونيـة للتدقيـق المتصـل بالانتخابـات تنتشـر فـي جميـع أنحـاء العالـم. ومـن الأمثلـة علـى ذلـك مبـادرة Verificado المكسـيكية فـي عـام 2018 [159] بقيـادة Animal Politico، وصحيفـة Newsweek باللُّغـة الإسـبانية، Pop Up Newsroom ومنصـة AJ + بالإسـبانية. وعملـت المبـادرة علـى كشـف «الأخبـار الزائفـة» والتدقيـق فـي الخطـاب السياسـي خـلال الحملـة الانتخابيـة المكسـيكية لعـام 2018. وكانـت رائـدةً مـن حيـث الحجـم؛ حيـث ضمـت أكثـر مـن 60 جهـةً مشـاركةً مـن الوسـائل الإعلاميـة ومنظمـات المجتمـع المدنـي والجامعـات، وجميعهـا يعمـل علـى مسـاعدة المواطنيـن علـى اتخـاذ قـرار الاقتـراع بنـاءً علـى معلومـات صحيحـة. وكان كل تقريـر يحمـل وسـم «Verificado_2018#» يعنـي أنـه قـد خضـع لمراجعـة شـبكة الشـركاء بأكملهـا وحظـي بدعمهـم. وقـام تمويـل Verificado 2018 مشـروع فيـس بـوك للصحافـة، ومبـادرة أخبـار Google الرقميـة، وتويتـر، فضـلاً عـن تنظيـم المكسـيكيين لمُكافحـة الفسـاد والإفـلات مـن العقـاب، ومؤسسـات أخـرى علـى غـرار المجتمـع المفتـوح ومنظمـة أوكسـفام. وفـازت المبـادرة بجائـزة الجمعيـة الأمريكيـة للصحافـة علـى الإنترنـت للتميُّـز فـي التعـاون والشـراكات لعـام 2018.

وكانـت Verificado 2018 مبـادرةً رائـدةً أيضـاً مـن حيـث الدعـم الـذي تلقتـه مـن شـركات التواصـل عبـر الإنترنـت. ومـن مفاتيـح نجاحهـا: 1. توفيـر الشـركات نفـاذاً إلـى بيانـات عـن أكثـر القصـص التـي تمـت مشـاركتها أو عـن أكثـر نتائـج البحـث رواجـاً، و2. تدابيـر تتعلـق بتنظيـم المُحتـوى للترويـج للمعلومـات التـي تـم التحقُّـق منهـا. والجديـر بالذكـر أن المنصـات لا تقـوم عـادةً بتزويـد وسـائل الإعـلام والجهـات المسـتقلة للتحقق مـن الوقائـع بهـذا الدعـم الشـامل أو بالنفـاذ إلـى بياناتهـا؛ ممّـا يعقـد عملهـا بشـكل كبيـر.

ومـن ميـزات مشـروع Verificado 2018 اعتمـاده وتكييفـه لمنصـة تعاونيـة أُنشِـأت أصـلاً علـى يـد Verificado19S لإدارة اسـتجابات المواطنيـن وعمليـات الإنقـاذ التعاونيـة.[160]

158 https://firstdraftnews.org/verification-toolbox/

159 https://verificado.mx/metodologia/

160 https://verificado19s.org/sobre-v19s/

وفي **أمريكا اللاتينية** برزت مبادرة Comprova [161]، وهي شراكة تضم 24 وسيلة إعلام برازيليةً، أُنشئت من أجل انتخابات عام 2018 ولكنها لا تزال مستمرة. ويتم تنسيق هذا المشروع من قبل Abraji (الرابطة البرازيلية للصحافة الاستقصائية) بالتعاون مع First Draft. وكما هو الحال في العديد من مشروعات التحقُّق من الوقائع التعاونية الأخرى، توفر مبادرة الأخبار من غوغل ومشروع الصحافة من فيس بوك الدعم المالي والتقني لـ Comprova. أما Projor، فهي منظمة غير هادفة للربح تركز في عملها على القضايا المتعلقة بوسائل الإعلام البرازيلية، كانت أيضاً من أوائل الداعمين للمشروع. وخلال الانتخابات، قام مشروع Comprova برصد المعلومات المنشورة من قِبل مصادر غير رسمية على وسائل التواصل الاجتماعي وتطبيقات المراسلة، كما وقام بالتحقُّق من صحتها (على تطبيق واتساب بشكل أساسي) (Tardáguila & Benevenuto et al., 2018). واعتمد رصد تطبيق واتساب على تقديم اقتراحات من الجمهور حول المُحتوى الذي يجب التحقُّق منه، فتم تقديم 67000 منشور. وهذا يدل بوضوح على الكم الهائل من المُحتوى السياسي الذي قد يكون إشكالياً والذي ينتشر على هذه التطبيقات المغلقة. كما ويشير إلى استحالة التدقيق والتحقُّق من وقائع هذا النوع من المُحتوى.

ومثال بارز آخر يأتي هذه المرة من الأرجنتين، وهو Reverso [162] وهو مشروع تعاوني ضخم، روّجت له ونسقته كل من منظمة Chequeado لتتحقق من الوقائع، وAFP Factual، وFirst Draft وغرفة أخبار Pop-Up. وقد ضم المشروع 100 جهة مشاركة من وسائل إعلام وشركات تكنولوجية اجتمعت خلال حملة الانتخابات الرئاسية الأرجنتينية لعام 2019. ولتحقيق التغطية الأوسع، قام مشروع Reverso بكشف زيف 180 مقالةً و30 مقطع فيديو طوال فترة الحملة التي استغرقت ستة أشهر. وعمد جميع الشركاء إلى نشر مخرجات المشروع في الوقت نفسه. كذلك قام فريق المشروع برصد فيس بوك وانستغرام وتويتر وتطبيقات المراسلة الخاصة (واتساب بشكل أساسي)؛ والمنصات على غرار يوتيوب وChequeo Colectivo [163] (وهي منصة تعبئة جماعية أو crowdsourcing تديرها Chequeado). أما من ناحية الابتكار ومن خلال التعاون مع BlackVox [164] تمكّن فريق Reverso أيضاً من التحقُّق من ملفات صوتية زائفة[165] للمرشحين؛ تم تناقلها على تطبيق واتساب[166].

والمثال البارز الأخير من أمريكا اللاتينية هو مشروع Verificado.UY في الأوروغواي[167] وهو يضم أكثر من 30 شريكاً قاموا برصد وكشف زيف المعلومات المُضلّلة خلال الانتخابات الرئاسية في أوروغواي في تشرين الأول/أكتوبر 2019 وقد حظي بدعم من منظمة First Draft لفائدة التدريب والدعم المالي والتكنولوجي. وركّز المشروع على نوعين من التحقُّق: التحقُّق من الشائعات التي تنتشر على الشبكات الاجتماعية، وتصريحات السياسيين والمرشحين.

161 https://projetocomprova.com.br/about/
162 https://reversoar.com/
163 https://chequeado.com/colectivo/
164 https://blackvox.com.ar/
165 https://www.clarin.com/politica/reverso-creo-nuevo-metodo-conicet-verificar-audios-virales-whatsapp_0_1638FLWL.html
166 https://www.poynter.org/fact-checking/2019/meet-forensia-a-software-ready-to-debunk-fakeواتساب-audio-files/
167 https://verificado.uy/

وفـي **أستراليا وآسيا** علـى التوالـي، تتضمّـن الأمثلـة مشـروع CrossCheck Australia [168] (الـذي تديـره First Draft)، والـذي راقـب الانتخابـات الفيدراليـة الأسـترالية فـي عـام 2019 ومشـروع Checkpoint التعاونـي فـي الهنـد الـذي أُطلق قبل الانتخابـات الوطنيـة فـي الهنـد فـي عـام 2019 (مراجعـة القسـم 6.1 لمزيـد مـن التفاصيـل).

وفـي **أفريقيا**، أقامـت First Draft شـراكةً مـع المركـز الدولـي للتقاريـر الصحفيـة الاسـتقصائية فـي نيجيريا و16 غرفـة أخبـار لإنشـاء CrossCheck Nigeria [169] فـي الفتـرة التـي سـبقت الانتخابـات النيجيريـة فـي شـباط/فبرايـر 2019. وقـدم الدعـم لهـذا المشـروع «مؤسسـة المجتمـع المفتـوح» أو the Open Society Foundation وقـد بُنـى المشـروع علـى المعرفـة والتكنولوجيـا مـن مبـادرات تعاونيـة سـابقة لمنظمـة First Draft، بمـا فـي ذلـك Comprova فـي البرازيـل وCrossCheck فـي فرنسـا. ومـن السـمات الرئيسـية لهـذا العمـل اعتمـاده منهجيـة «CrossCheck» التـي تقضـي بقيـام صحفييـن مـن مختلـف غـرف الأخبـار بالتدقيـق فـي عمـل بعضهـم البعـض لضمـان الالتـزام بمبـادئ الشـفافية والدقـة وعـدم التحيـز.

مثـال آخـر هـو مشـروع Real411 فـي جنـوب أفريقيـا[170] (علمـاً أنّ الرمـز العـددي ‹411› يعنـي «المعلومـات» باللُّغـة المسـتخدمة علـى الإنترنـت). وعلـى عكـس المشـروعات السـابقة التـي أطلقتهـا وسـائل الإعـلام ومنظمـة First Draft فـإنّ مشـروع Real411 أطلقتـه منظمـة غيـر حكوميـة (Media Monitoring Africa)، وسـاهمت فيـه اللجنـة الانتخابيـة فـي جنـوب أفريقيـا. وعلـى غِـرار المبـادرات الأخـرى، فإنـه يوفّـر منصـةً علـى الإنترنـت تسـمح للمواطنيـن بالتبليـغ عـن المعلومـات المُضلّلـة المزعومـة. ولكـن هـذه المنصـة تشـتمل علـى اسـتجابة حكوميـة حيـث أنهـا موصولـة بمديريـة الجرائـم الانتخابيـة وتنظـر فـي الشـكاوى لجنـة تضـم خبـراء بمـا فـي ذلـك فـي قانـون الإعـلام، كمـا وتضـم ممثليـن عـن وسـائل الإعـلام الاجتماعيـة والرقميـة. وترفـع هـذه اللجنـة توصياتهـا للمفوضيـة العليـا للانتخابـات حـول إمكانيـة اتخـاذ مزيـد مـن الإجـراءات، بمـا فـي ذلـك إحالـة القضايـا للقضـاء لكـي يصـار إلـى اتخـاذ الإجـراء القانونـي الجنائـي أو المدنـي المناسـب؛ والطلـب مـن منصـات التواصـل الاجتماعـي إزالـة المـواد المسـيئة وإصـدار التصريحـات الإعلاميـة لتنبيـه الجمهـور وتصحيـح المعلومـات المغلوطـة. ويحتـوي موقـع Real411 علـى قاعـدة بيانـات تضـم جميـع الشـكاوى الـواردة ومسـتجداتها. وللمسـاعدة فـي التمييـز بيـن الإعلانـات الرسـمية وتلـك الزائفـة؛ طُلـب مـن الأحـزاب السياسـية المشـاركة فـي انتخابـات 8 أيار/مايـو 2019 تحميـل جميـع المـواد الإعلانيـة الرسـمية المسـتخدمة مـن قِبلهـا علـى مخـزن رقمـي للإعلانـات السياسـية علـى الإنترنـت علـى الموقـع التالـي: www.padre.org.za. كمـا وتـم تكييـف هـذه المبـادرة للتصـدي للمعلومـات المُضلّلـة حـول جائحـة فيـروس كورونـا.

الاستجابات الدولية التعاونية

الاسـتجابات التعاونيـة علـى مسـتوى الاتحـاد الأوروبـي: تأسـس مشـروع FactCheckEU.info [171] مـن قِبـل الشـبكة الدوليـة للتحقـق مـن الوقائـع (IFCN)، وهـو يضـم 19 وسـيلةً إعلاميـةً أوروبيـةً مـن 13 دولـةً (هـي الـدول الأوروبيـة الموقعـة علـى مدونـة مبـادئ الشـبكة) وكان يهـدف إلـى مُكافحـة التّضليـل فـي الاتحـاد الأوروبـي قبـل

[168] https://firstdraftnews.org/project/crosscheck-australia/
[169] https://crosschecknigeria.org/about/faqs
[170] https://www.real411.org/
[171] https://factcheckeu.info/en/

انتخابات البرلمان الأوروبي أيار/مايو 2019 (Darmanin, 2019). ويركز المشروع بشكل أساسي على كشف زيف المعلومات المُضلّلة ونشر الحقائق عن أوروبا للحد من المفاهيم الخاطئة (مثل: معاداة الإسلام والهجرة). ويمكن للمواطنين من خلال هذا المشروع تقديم مطالبات للتحقق من معلومات معيّنة من خلال استمارة أسئلة وأجوبة على شبكة الإنترنت يتم الرد عليها من قبل أحد الشركاء. ولتحقيق التغطية القصوى، نُشرت المقالات بلغاتها الأصلية وتمت ترجمتها إلى اللُّغة الإنكليزية.

وكانت هذه المبادرة مستقلةً تماماً عن مؤسسات الاتحاد الأوروبي وغيرها من الجهات الفاعلة الحكومية. وقد قامت صحيفة Libération ووكالة الإنترنت Datagif بإنشاء المنصة من خلال منحة ابتكار قدرها 50000 دولار أميركي من معهد بوينتر. وقدم غوغل 44000 يورو، ومبادرة المجتمع المفتوح لأوروبا 40000 يورو، والشبكة الدولية للتحقق من الوقائع 10000 يورو لتغطية التكاليف الأخرى، أي راتب منسق المشروع بدوام كامل لمدة ستة أشهر، وتكاليف ترجمة المُحتوى.

وبالتوازي، أُطلقت مبادرتان تعاونيتان أُخريان للتحقق من الوقائع وهما: مرصد التّضليل الذي يموله الاتحاد الأوروبي (وهو SOMA الذي تم استعراضه في القسم 4.1)[172]، وCrossCheck أوروبا التي أطلقتها First Draft[173].

وعند كتابة هذا التقرير البحثي، كانت مبادرة CrossCheck من First Draft تتوسع كشبكة عالمية من الصحفيين والباحثين يعملون بشكل تعاوني للتحقيق في المُحتوى الموجود على الإنترنت خلال الانتخابات وبعدها. وبناءً على الحملات السابقة في الولايات المتحدة وفرنسا والبرازيل ونيجيريا وإسبانيا وأُستراليا والاتحاد الأوروبي، تسعى هذه المبادرة إلى إثبات قدرة المؤسسات الإخبارية على العمل معاً على نطاق عالمي لكتابة تقارير أكثر فعاليةً وكفاءةً ومسؤوليةً لمُكافحة التّضليل.

هـ. استجابات شركات التواصل على الإنترنت:

قبل الانتخابات الرئاسية الأمريكية في عام 2020، كانت المخاوف تتصاعد حول تفاقم الاستقطاب، والتدخل الأجنبي، وظهور أشكال جديدة من التلاعب بالمُحتوى الرقمي على غِرار ما يُسمّى بالتزييف الفج أو deepfakes وهي مقاطع فيديو أو ملفات صوتية ملفقة يتم إنتاجها من خلال التعلم الآلي (يُرجَى مراجعة القسم 6.2 من الفصل السادس). على هذه الخلفية، أعلن نائب رئيس فيس بوك للتواصل والشؤون العالمية؛ نائب رئيس الوزراء البريطاني الأسبق «نك كليغ» في عام 2019 أن «فيس بوك اقترف أخطاءً في عام 2016» لكنه أضاف أن الشركة أمضت السنوات الثلاث بعد ذلك على «بناء دفاعاتها لمنع تكرار ذلك مرةً أخرى» (Clegg, 2019). ثم عدَّدَ الإجراءات التي اتخذها فيس بوك للقضاء على الحسابات «الزائفة» – والتي وصفها بأنها مصادر رئيسية «للأخبار الزائفة» والمحتويات الخبيثة. فقد استقدم فيس بوك مثلاً «جهات مستقلةً للتحقق من صحة المُحتوى» (يُرجَى مراجعة القسم 4.1 لتحليل شبكة التحقُّق من الوقائع عبر أطراف ثالثة التابعة لفيس بوك والقسم 7.1 لتقييم القضايا الأخلاقية ذات الصلة) كما قام «بتجنيد جيش من الأفراد– يبلغ اليوم نحو 30000 شخص والاستثمار بشكل كبير في أنظمة الذكاء الاصطناعي لإزالة المحتويات الضارة».

....

172 https://www.disinfobservatory.org/the-observatory/

173 https://firstdraftnews.org/project/crosscheck-europe/

وفي ما يخص الإعلانات السياسية المغلوطة أو المُضلّلة، فقد تعرّض فيس بوك لانتقادات شديدة بسبب سياسته في هذه المجال. فقد اتهمت السيناتورة «إليزابيث وارن» فيس بوك بتحويل منصته «إلى آلة تضليل هادفة للربح». ولإثبات ادعائها نشرت إعلاناً زائفاً هو التالي «خبر عاجل: مارك زوكربيرغ وفيس بوك يؤيدان إعادة انتخاب دونالد ترامب لولاية ثانية.»[174] وأفيد في تشرين الأول/أكتوبر 2019 أنّ فيس بوك قد عدّل قواعده المتعلقة بمنع الإعلانات ذات المُحتوى «الخاطئ والمضلل». فبعد أن كان ذلك يُعرّف بأنه «أي مُحتوى خادع، خاطئ أو مضلّل بما في ذلك الادعاءات، والعروض، والطرق المخادعة»، أصبح هذا التعريف أضيق ليتم حظر «الإعلانات التي تتضمن ادعاءات تم كشف زيفها من قبل أطراف ثالثة تعمل في مجال التحقُّق من الوقائع أو، في ظروف معينة، من قبل منظمات ذات خبرة خاصة.» (Legum, 2019).

ويضع فيس بوك قيوداً على التحقُّق من السياسيين والأحزاب السياسية من خلال مبادئ توجيهية موجهة لشركائه العاملين في مجال التحقُّق من الوقائع والتي تنص على التالي: «عموماً، لا تخضع منشورات وإعلانات السياسيين إلى التحقُّق» (Facebook, 2019b). وتتماشى المبادئ التوجيهية مع «إيمان فيس بوك الراسخ بحُريّة التّعبير، واحترام العملية الديمقراطية، وإيمانه بأنه خاصةً في الديمقراطيات الناضجة التي تتمتع بصحافة حرة يُعتبر الخطاب السياسي من أكثر الخطابات التي تخضع للتدقيق». (مراجعة القسمين 4.1 و7.1 لمزيد من التفاصيل حول هذه القضايا). وتشير المبادئ التوجيهية إلى أنه «إذا قام السياسي بادعاء مباشرةً على صفحته أو في أحد الإعلانات أو على موقعه على الإنترنت، يُعتبر هذا الادعاء خطاباً مباشراً لا يخضع لبرنامج التحقُّق من الوقائع من قبل أطراف ثالثة حتى ولو كان مضمون الادعاء قد تم كشف زيفه في مكان آخر.»

وعلى نقيض مقاربة فيس بوك القاضية بعدم التدخل في التّضليل السياسي، أعلن الرئيس التنفيذي لشركة تويتر «جاك دورسي» أن المنصة التي أسسها –تويتر– ستتوقف عن نشر كافة الإعلانات السياسية بدءً من 22 تشرين الثاني/نوفمبر 2019. وقال أنّ هذا يعكس المخاوف من أن «يؤدي دفع الأموال من أجل زيادة انتشار الخطاب السياسي إلى تداعيات كبيرة قد لا تكون البنية التحتية الديمقراطية اليوم مستعدةً لمواجهتها»[175]. وكما نوقش سابقاً في هذا الفصل، طرأ خلاف علني بين تويتر وفيس بوك في منتصف عام 2020 حول التحقُّق من الوقائع وكشف زيف المحتويات المنشورة من قِبل السياسيين، في حادث نجم عن قرار تويتر، للمرة الأولى، وضع تحذير على التغريدات المُضلّلة للرئيس الأمريكي في ما يخص العمليات الانتخابية.

وفي كانون الأول/ديسمبر 2019، أعلنت شركة غوغل الالتزام باتخاذ «مجموعة واسعة من التدابير للمساعدة في حماية الحملات وإبراز الأخبار الموثوقة حول الانتخابات وحماية الانتخابات من التدخل الأجنبي» (Spencer, 2019). وأكدت الشركة على أنها تريد «تحسين ثقة الناخبين في الإعلانات السياسية التي قد تظهر على منصات الإعلان لدينا». كما وأعلنت عن تغييرات بما في ذلك الحد من الإعلانات الانتخابية والاستهداف الدقيق للجمهور وفقاً للسن والنوع الاجتماعي والموقع العام (مستوى الرمز البريدي). وأوضحت أيضاً سياستها المتعلقة بالإعلان من خلال حظر «المقاطع ذات الزيف العميق أو الفج»، والادعاءات المُضلّلة حول عملية التعداد، والإعلانات التي يَثْبُت أنها تنشر ادعاءات كاذبة من شأنها تقويض المشاركة أو الثقة في العملية الانتخابية أو الديمقراطية بشكل كبير.

....

174 https://twitter.com/ewarren/status/1183019880867680256

175 https://twitter.com/i/events/1189643849385177088

وكجزء من الشفافية المتبعة في ما يخص الإعلانات الانتخابية على غوغل، صرّحت الشركة أنها توفّر إفصاحات في الإعلانات وتقارير شفافية على الإنترنت[176] تُظهر المُحتوى الفعلي للإعلان، ومن دفع لوضع الإعلان وكم أنفق وعدد الأشخاص الذين شاهدوا الإعلان وكيف تم استهدافهم (وهذا متوفر فقط بالنسبة إلى أوروبا والمملكة المتحدة والهند والولايات المتحدة). وصرّحت الشركة أنها «تتوقع أن تكون الإعلانات السياسية التي ستتطلب اتخاذ إجراءات بشأنها محدودةً للغاية، ولكن سوف نستمر في القيام بذلك بشأن الانتهاكات الواضحة». ومع ذلك، واجهت الشركة انتقادات شديدة قبل الانتخابات الأمريكية عام 2020 عندما رفضت إزالة إعلانات مجموعة متهمة بقمع مشاركة الناخبين، وذلك لادعاءاتها كذباً أن هناك فرق جوهري بين التصويت الغيابي والتصويت عن طريق البريد. ومع ذلك، وافق فيس بوك على إزالة إعلانات مشابهة من المجموعة نفسها (Stanley-Becker 2020).

و. الاستجابات التنظيمية للتضليل الانتخابي:

تلعب اللجان الانتخابية أو الوحدات الحكومية المتخصصة أيضاً دوراً رئيسياً في محاربة التّضليل الانتخابي من خلال استجابات مستهدفة. ومن الأمثلة على ذلك الإجراءات التي اتخذتها اللجنة الانتخابية الأُسترالية في 2019، بما في ذلك الإذن بالتواصل الانتخابي (AEC, 2019a)، والوحدة الحكومية الإسبانية الخاصة «بالتهديدات المختلطة» التي تركز على الأمن والرصد وأحياناً دحض المعلومات المُضلّلة على الإنترنت (Abellán, 2019).

وبطبيعة الحال، يمكن ضمان سلامة الانتخابات من خلال تدابير تشريعية. وتتم مناقشة ذلك بشكل محدد في القسم 5.1 والملحق أ، في القسمين الأول حول التشريعات المقترحة والثاني حول التشريعات المعتمدة.

كما وبإمكان اللجان الانتخابية واللجان الحكومية أيضاً أن توفّر معلومات موثوقة عن المرشحين والأحزاب، والعمل مع شركات التواصل عبر الإنترنت من أجل تعزيز هذه المعلومات. فعلى سبيل المثال، أنشأت الحكومة الكندية بروتوكول الحوادث الانتخابية الخطيرة في عام 2019 [177] كآلية لإبلاغ المواطنين عن المخاطر التي تهدد سلامة الانتخابات وإعلام المرشحين والمنظمات ومسؤولي الانتخابات المستهدفين بالهجمات. مثال آخر، الوزارة الإندونيسية للاتصالات وتكنولوجيا المعلومات التي أنشأت في عام 2019 «غرفة عمليات» لكشف وتعطيل المحتويات السلبية والانتهاكية (Board, 2019). ومن الأمثلة على الاستجابات التعاونية «مبادرة المعهد المكسيكي الوطني للانتخابات» (INE)، الذي وقّع اتفاقية تعاون مع كل من فيس بوك وتويتر وغوغل للحد من انتشار التّضليل الانتخابي، ولنشر المعلومات الانتخابية العملية خلال انتخابات العامين 2018 و2019.

ومن أنواع الاستجابات التنظيمية المهمة الأخرى، الاستجابات التي تستهدف شفافية ونزاهة الإعلانات على الإنترنت خلال الفترات الانتخابية. فعلى سبيل المثال، في عام 2019 تقدمت الحكومة الآيرلندية باقتراح تشريعي يهدف إلى تنظيم شفافية الإعلانات السياسية المدفوعة على الإنترنت خلال الفترات الانتخابية (Irish Department of the Taoiseach, 2019). وتقضي المقاربات المكمّلة بتشجيع الأحزاب السياسية على تسجيل إعلاناتها على شبكة الإنترنت في قاعدة بيانات متاحة للجمهور أو اعتماد تشريع

176 https://transparencyreport.google.com/political-ads/home?hl=en

177 https://www.canada.ca/en/democratic-institutions/services/protecting-democracy/critical-electionincident-public-protocol.html

يفرض عليها ذلك. ففي 2019، قامت المفوضية العليا للانتخابات في جنوب أفريقيا على سبيل المثال بإنشاء مخزن رقمي للإعلانات السياسية للأحزاب لهذا الغرض على الموقع التالي padre.org.za.

وقد قامت الاستجابات بإشراك المواطنين لمساعدتها على اكتشاف المعلومات الانتخابية المُضلّلة والتبليغ عنها واتخاذ إجراءات جادة بشأنها. ومن الأمثلة على ذلك، كما سبق وذُكر أعلاه، بوابة real411.org التي تم إنشاؤها بالتعاون مع المفوضية العليا للانتخابات في جنوب أفريقيا. مثال آخر، بوابة «الزر الأحمر» التابعة للحكومة الإيطالية، والتي تسمح للمواطنين بتبليغ وحدة الشرطة الخاصة بالإنترنت بالمعلومات المُضلّلة. فتقوم وحدة الشرطة بالتحقيق في المُحتوى، كما وتساعد المواطنين على تبليغ شركات التواصل عبر الإنترنت بالمعلومات المُضلّلة، ورفع دعوى قضائية في حال كانت المحتويات تشهيريةً أو غير قانونية (la Cour, 2019).

ومن الاستجابات الأخرى إغلاق الإنترنت، على الرغم من أن هذا التدبير يُعتبر غير متناسب وقد تكون له نتائج عكسية على المصداقية خلال الانتخابات. وقد فرضت بعض الحكومات هذه الاستجابات في الفترات السابقة للانتخابات بحجة أن هدفها حماية المواطنين من التّضليل والدعاية الانتخابيين (Al Jazeera, 2018; Paul, 2018).

وهناك أيضاً بعض الأمثلة على استجابات دولية. فقد اعتمد الاتحاد الأوروبي خطة عمل بشأن التّضليل، قبل الانتخابات الأوروبية لعام 2019، تهدف إلى بناء القدرات والتعاون داخل الاتحاد الأوروبي وبين الدول الأعضاء (Al Jazeera, 2018; Paul, 2018). كما وتُدير دائرة العمل الخارجي الأوروبي موقعاً على شبكة الإنترنت يهدف إلى توفير روايات مضادة للمعلومات المُضلّلة. ومثال آخر هو الدليل الذي نشرته منظمة الدول الأمريكية لضمان حُريّة التّعبير في ما يخص التّضليل المتعمد في السياقات الانتخابية (OAS, 2019). وهو يقدّم توصيات لمجموعة واسعة من الجهات الفاعلة، منها: السلطة التشريعية، والسلطة القضائية، والسلطة التنفيذية، والسُّلطات الانتخابية، وشركات التواصل عبر الإنترنت، والأحزاب السياسية، وشركات الاتصالات، ووسائل الإعلام والصحفيين، وجهات التحقُّق من الوقائع، والشركات التي تتداول البيانات التجارية لأغراض الدعاية، والجامعات ومراكز البحوث (OAS, 2019).

5.3.5 كيف يتم تقييم الاستجابات الانتخابية؟

بما أنّ العديد من الاستجابات الانتخابية تم تغطيتها ضمن فئات الاستجابات الأخرى (مثل: استجابات التحقُّق من الوقائع، والاستجابات التنظيمية) التي تُستخدم بشكل محدد لاستهداف المعلومات الخاصة بالانتخابات، فإنّ الأساليب والنتائج المستخلصة من التقييمات الخاصة بكل من هذه الأنواع تطبّق أيضاً على الاستجابات الانتخابية.

وفي ما يتعلق بشفافية الإعلانات السياسية، يؤكد فيس بوك على أنّ «الإعلانات حول القضايا الاجتماعية والانتخابات أو السياسة ملزمة بمستوى أعلى من الشفافية». ويضيف: «سيتم وضع كافة الإعلانات النشطة وغير النشطة للسياسيين على فيس بوك في مكتبة علنية[178] متاحة للجمهور وقابلة للبحث لمدة تصل إلى سبع سنوات»، مما يسمح بالتقييم.

[178] https://www.facebook.com/ads/library/

ومع ذلك، خلُصَت دراسة «معهد الحوار الاستراتيجي» (ISD, 2019) حول الانتخابات الأوروبية، أنّ مكتبة فيس بوك للإعلانات «تشوبها أوجه قصور متعددة». فتصنيفها للإعلانات هو «غالباً اعتباطي». فعلى سبيل المثال تم اتهام هذه المكتبة بالأخطاء بتصنيف «تجمّع مهندسي التدفئة» في إيطاليا ولعبة الحاسوب «زنزانات وتنانين» في ألمانيا حيث صنفتهما على أنهما من المحتويات السياسية، في حين أنّ إعلانات الحزب الألماني اليميني المتطرّف لم توضع في مكتبة الإعلانات. وفي إحدى التدوينات[179] اشتكت «موزيلا» أيضاً أنّ واجهة برمجة أرشيف الإعلانات التابع لفيس بوك «غير ملائمة»، حيث أنها لم تستوف إلا معيارين من معايير الخبرة الدُّنيا.

وتؤكد الدراسة نفسها (ISD, 2019) أن شركات التواصل عبر الإنترنت «تضع حُريّة التّعبير في خطر، عبر فرض رقابة مفرطة ومضلَلة، في حين أنها غير قادرة على مواكبة العديد من الحملات والتكتيكات الخبيثة»، علماً أنّ هذه الأخيرة تمثل أيضاً خطراً على حُريّة التّعبير. كما وأشارت الدراسة إلى اتخاذ تدابير أدت إلى نتائج عكسية في ألمانيا، على سبيل المثال، حيث استغلت شبكات اليمين المتطرّف محاولات تويتر تسريع عملية التبليغ عن المعلومات المُضلّلة، ممّا أدى إلى إزالة أو تعليق حسابات نشطاء مناهضين لليمين المتطرف وصحف تمثل المصالح اليهودية، كما وتمت إزالة أو تعليق حسابات ضحايا المضايقات بدلاً من حسابات مرتكبي هذه المضايقات.

6.3.5 التحديات والفرص

تشير البحوث التي أُجريت مؤخراً إلى تطوّر أساليب التّضليل. ويتوقّع معهد الحوار الاستراتيجي السابق ذكره أن «تكون الأحزاب الشعبوية واللّجان الإلكترونية لليمين المتطرف والجماعات الدينية قادرةً على التكيّف مع التكتيكات المستخدمة أكثر من الدول.» ويشير المعهد، ومقره لندن، إلى تطوّر ملحوظ وهو «الابتعاد عما يسمى «الأخبار الزائفة» واعتماد «السرديات المتنافسة»، مع تعزيز حركيّة «صراع الثقافات» حول قضايا على غرار الهجرة، والمسلمين في أوروبا، والقيم الأسرية مقابل القيم التقدمية، وبشكل متزايد، السياسة المناخية» (ISD, 2019). والنتيجة هي أن الصلة بين الأحزاب السياسية والمُحتوى على الإنترنت تصبح في معظم الأحيان غير واضحة المعالم أو مبهمةً تماماً. كذلك يصبح بالإمكان إخفاء هُويّات الأفراد الكامنين وراء الرسائل ولا تُعتمد الشفافية حول الآليات المعتمدة لتحقيق «الوصول».

وفي حملة انتخابات الاتحاد الأوروبي لعام 2019، استخدمت شبكة Avaaz غير الربحية (Avaaz, 2019) منصة التعبئة الجماعية[180] لتحديد التكتيكات الجديدة للشبكات اليمينية المتطرفة في مختلف أنحاء الاتحاد الأوروبي والتي بدأت باعتماد الممارسات التالية: استخدام حسابات وهمية ومنسوخة لتضخيم انتشار التّضليل، تنسيق غير طبيعي بين وسائل بديلة محددة لمشاركة مُحتوى وخطاب الكراهية نفسه وإعادة تدوير المتابعين عبر تغيير مضلل لأسماء الصفحات؛ استخدام العناوين المثيرة للانتباه أو clickbait، والترويج لأجندات سياسية أو انقسامية من خلال صفحات ترفيه تتمتع بالانتشار. ويتمثل التحدي في السماح بهذا الرصد والتعرض خلال الانتخابات، وعلى نطاق واسع.

....

179 https://blog.mozilla.org/blog/2019/04/29/facebooks-ad-archive-api-is-inadequate/

180 https://fake-watch.eu/

فـي تقرير حـول الانتخابـات العامـة فـي المملكـة المتحـدة لعـام 2019 (Election Monitoring, 2019)، أشـارت ثمانـي منظمـات إلـى انعـدام الشـفافية فـي قيـام الأحـزاب السياسـية بجمـع ومعالجـة بيانـات الناخبين، كما وإلى انعـدام الشـفافية فـي الإعلانـات السياسـية والرسـائل الهادفـة، بمـا فـي ذلـك ادعـاءات مبالـغ فيهـا ومضللـة. وأشـارت مخـاوف إضافيـة إلـى «ترتيبـات تمويـل غامضـة» تهـدف إلـى الدفـع «بالمُحتـوى المدفـوع» علـى الناخبيـن، وإلـى نشـاطات مشـابهة لنشـاط روبوتـات الإنترنـت فـي المناقشـات حـول الأحـزاب السياسـية والسياسـات، وقيـام الجهـات الفاعلـة شـديدة التحـزب بنشـر التّضليـل ونظريـات المؤامـرة علـى فيـس بـوك، ومضايقـة شـخصيات سياسـية رئيسـية وصحفييـن علـى الإنترنـت، وحتـى إنشـاء منظمـات اقتـراع متحيـزة. ودعـا الموقعـون الثمانيـة، بمـا فـي ذلـك معهـد الحـوار الاسـتراتيجي، ومنظمـة Full Fact ومشـروع الدعايـة الحاسـوبية فـي جامعـة أكسـفورد، إلـى إصلاحـات انتخابيـة لمواجهـة هـذه المخاطـر الرقميـة التـي تهـدد الديمقراطيـة.

ووُجهـت دعـوات إلـى شـركات التواصـل عبـر الإنترنـت لإبـراز مصـادر المعلومـات الموثـوق بهـا وترويجهـا خـلال الانتخابـات، ولمُكافحـة مسـائل علـى غِـرار خـداع الناخبيـن (علـى سـبيل المثـال، الحـد مـن نسـبة اقتـراع الناخبيـن) أو تقويـض الثقـة فـي العمليـة الانتخابيـة (Goldzweig, 2020). وكجـزء مـن اسـتجاباتها للمعلومـات المُضلّلـة حـول جائحـة فيـروس كورونـا، أثبتـت الشـركات أنهـا تتمتـع بالقـدرات التقنيـة للقيـام بذلـك (Posetti & Bontcheva, 2020a; Posetti & Bontcheva, 2020b) ويكمـن التحـدي بالنسـبة إليهـا فـي تكييـف هـذه القـدرات بهـدف الترويـج للمعلومـات التـي يمكـن الاعتمـاد عليهـا مـن مصـادر موثوقـة خـلال الانتخابـات، علـى غـرار الهيئـات الانتخابيـة و/أو الهيئـات المسـتقلة لمراقبـة نزاهـة الانتخابـات.

ومـن التحديـات الأخـرى التـي لابـد مـن معالجتهـا، نمـوذج تمويـل هيئـات التدقيـق والتحقُّـق مـن الوقائـع والشـفافية المحـدودة المرتبطـة بهـذه الجهـود أحيانـاً. علـى سـبيل المثـال، إذا كانـت شـركة مـا للتواصـل عبـر الإنترنـت تتحكّـم فـي معاييـر عمليـات التحقُّـق التـي تتـم علـى مواقعهـا مـن قبـل أطـراف ثالثـة خـلال الانتخابـات، ورفضـت تمويـل التحقُّـق مـن مُحتـوى معيـن أو تطبيـق نتائـج هـذا التحقُّـق، مـن الطبيعـي أن يؤثـر ذلـك علـى فعاليـة عمليـة التحقُّـق مـن الوقائـع والنظـرة إلـى اسـتقلالية وموثوقيـة جهـود التحقُّـق هـذه.

وبالمثـل، إذا كانـت هـذه الشـركات تقـوم بتمويـل الهيئـات غيـر الربحيـة للتحقـق مـن الوقائـع ومعاهـد البحـوث التـي تحقـق فـي المحتويـات المُضلّلـة والشـبكات التـي تنتشـر علـى وسـائل التواصـل الاجتماعـي خـلال الانتخابـات، فمـا هـي التداعيـات المترتبـة علـى اسـتقلاليتها ومـا هـي الضمانـات التـي تؤكـد عـدم قيـام المموليـن بممارسـة ضغـوط غيـر مبـررة علـى هـذه الهيئـات؟

وتكتسـي هـذه الاعتبـارات أهميـةً خاصـةً فـي ضـوء التحـدي الكبيـر المطـروح علـى شـركات التواصـل عبـر الإنترنـت والمتمثـل فـي إقامـة التـوازن بيـن مسـؤولياتها فـي احتـرام حُريّـة التّعبيـر، والقيـام فـي الوقـت نفسـه وبشـكل مسـتمر خـلال الانتخابـات بالتحذيـر مـن المعلومـات المُضلّلـة والحـد منهـا وحظرهـا فيمـا تواجـه ضغوطـاً متزايـدةً مـن فاعليـن سياسـيين نافذيـن لدفـع هـذه الشـركات علـى معاملتهـم علـى أنهـم اسـتثناء لهـذه القواعـد.

ومعـاً، تسـلّط هـذه الأمثلـة الضـوء علـى التحديـات الكبيـرة التـي تحيـط بالتّضليـل خـلال الانتخابـات واسـتهداف الناخبيـن والتلاعـب بهـم. ومـع تنظيـم انتخابـات وطنيـة متعـددة فـي العالـم كل عـام، ومئـات الانتخابـات الإقليميـة والمحليـة؛ يُشـكّل هـذا تحـد كبيـر مسـتمر أمـام منصـات التواصـل علـى الإنترنـت والحكومـات فـي جميـع أنحـاء العالـم علـى حـد سـواء. ولكنـه فـي الوقـت نفسـه يتيـح فرصـاً كبيـرةً ويشـكل زخمـاً هائـلاً لجهـات التحقُّـق مـن

الوقائع المستقلة والصحفيين ووسائل الإعلام والمجتمع المدني والباحثين والمنظمات الوطنية والدولية يدفعها إلى مواصلة -بل وتوسيع- الدور الأساسي الذي تلعبه في مجال رصد المعلومات المُضلّلة وكشفها والتصدي لها والحد منها، وتقييم أثرها.

7.3.5 توصيات تتعلّق بالاستجابات الانتخابية

نظراً للتحديات والفرص المحددة أعلاه، والأضرار الكبيرة المحتملة للمعلومات المُضلّلة المصاحبة للانتخابات، يمكن تقديم التوصيات السياستية التالية.

توصيات للحكومات والمنظمات الدولية:

- الاستثمار في رصد وقياس وتقييم فعالية الاستجابات الانتخابية للمعلومات المُضلّلة.
- العمل مع شركات التواصل عبر الإنترنت لضمان أن تكون الاستجابات التي تعتمدها شفافةً وقابلةً للقياس بشكل مناسب، وأن تُطبَّق على نطاق عالمي.
- تشجيع منصات التواصل عبر الإنترنت على تطبيق استجابات حاسمة وسريعة للتصدي للتضليل الانتخابي، كما فعلت للتصدي للمعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا.
- تنسيق مبادرة تهدف إلى تعزيز الحفاظ على الخصوصية واحترامها، والنفاذ المتساوي إلى البيانات الرئيسية من منصات التواصل عبر الإنترنت، للسماح بكتابة البحوث المستقلة على نطاق جغرافي ملائم حول تواتر وانتشار وتأثير التّضليل على الإنترنت على المواطنين خلال الانتخابات.
- تسهيل وتشجيع التعاون بين أصحاب المصلحة المُتعددين على المُستوى العالمي، وتبادل أفضل الممارسات عبر القارات والدول، لتحقيق التنفيذ الفعال للتدابير الشاملة والهادفة إلى مواجهة التّضليل على الإنترنت خلال الانتخابات.

توصيات لمنصات التواصل عبر الإنترنت:

- الإقرار بالضرر الكبير المحتمل الناجم عن التّضليل السياسي، وتحديداً في الفترة التي تسبق الانتخابات (بما فيه المعلومات المُضلّلة في الإعلانات على شبكة الإنترنت)، والانخراط في حوار بين مختلف أصحاب المصلحة حول السياسات والأساليب المعتمدة خاصةً خلال الفترات الانتخابية. ويمكن أن يشمل ذلك قيوداً مؤقتةً على الإعلانات السياسية قبل الانتخابات، وتقديم مزيد من المعلومات لضمان الشفافية في ما يخص الإعلانات السياسية خلال الفترات الانتخابية، واعتماد سياسات انتخابية محددة للترويج لمصادر المعلومات الموثوقة، وتعزيز الجهود الآيلة إلى تنظيم المُحتوى وموارد التحقُّق من الوقائع.
- التصدي للمعلومات الانتخابية المُضلّلة العابرة للمنصات، والتعاون على وضع معايير على مستوى القطاع الأوسع للتصدي للتضليل الانتخابي ولدعم الديمقراطية والتنظيم الذاتي.
- التعاون على تحسين قدراتها في مجال كشف المعلومات المُضلّلة والحد منها، حيث أنّ مُمارسات التّضليل خلال الانتخابات غالباً ما تمارس بطرق عابرة للمنصات.

- تطبيق الدروس المستفادة من الاستجابة الطارئة لجائحة المعلومات المُضلّلة حول فيروس كورونا وتطبيق تلك الدروس لإدارة التّضليل السياسي والانتخابي.

- الإسهام بشكل فعال في تمويل الأبحاث المستقلة حول مظاهر وأثر التّضليل الانتخابي، والتقييم المستقل لفعالية استجابات الشركات للتضليل، على أن تموّل هذه المبادرات من قبل مجالس تمويل مستقلة تماماً.

- العمل معاً، وبتوجيه من مقرر الأمم المتحدة الخاص؛ المعني بتعزيز وحماية الحق في حُريّة الرأي والتّعبير، ومع خبراء دوليين مستقلين آخرين، لوضع مقاربة سياسية مُخطَّط لها، بحيث تهدف إلى التصدي لوكلاء التّضليل الذين يستخدمون مواقع المنصات على الإنترنت وهم لهم مناصب سياسية نافذة.

توصيات للهيئات الانتخابية الناظمة وللسلطات الوطنية:

- تعزيز التشريعات التي تساعد على حماية المواطنين من التّضليل الانتخابي (مثل: حماية البيانات وحُريّة التّعبير وشفافية الإعلانات الانتخابية).

- تحسين شفافية كافة الإعلانات الانتخابية التي تصدر عن الأحزاب السياسية والمرشحين من خلال إلزامهم بإنشاء قواعد بيانات شاملة ومتاحة علناً تتضمن إعلاناتهم والإفصاح عن الإنفاق على الإعلان من قبل الأحزاب السياسية وجماعات الدعم.

- إقامة تعاون فعّال مع شركات التواصل عبر الإنترنت لرصد ومعالجة الأخطار التي تُهدّد سلامة الانتخابات.

- إقامة وتعزيز الاستجابات متعددة الأطراف بين أصحاب المصلحة بما في ذلك المجتمع المدني بشكل خاص.

- العمل على تثقيف وتمكين المواطنين لكي يصبحوا قادرين على كشف المعلومات المُضلّلة خلال الانتخابات والتبليغ عنها.

- تحسين معرفة المواطنين بالعملية الانتخابية وانخراطهم فيها من خلال مبادرات التربية المدنية ومبادرات تعزيز الدراية.

- التعاون مع الوكالات الإخبارية والباحثين المتخصصين لإبراز المعلومات المُضلّلة والتحقيق في شبكات التّضليل.

توصيات لوسائل الإعلام ومنظمات التحقُّق من الوقائع المستقلة:

- النظر في توسيع التحقُّق من الوقائع خلال الانتخابات ليشمل البث الحي بصفةٍ عامّة، والبث الحي على الإنترنت على وجه الخصوص، لتعزيز نطاق الوصول والأثر.

- إجراء البحوث الهادفة إلى تقييم فعالية المقاربات المختلفة لكشف الزيف واحتواء التّضليل خلال الانتخابات، بما في ذلك الاستجابات التي تنفذها الهيئات التنظيمية ومنصّات التواصل عبر الإنترنت.

6

الفصل السادس:
الاستجابات المُندرجة
في إطار الإنتاج والتّوزيع



1.6 الاستجابات المتعلقة بتنظيم المُحتوى

المؤلفون: تريشا ماير، كلارا هانو، جولي بوزيتي

يُناقش هذا الفصل تدابير التّصدي للتضليل من خلال تنظيم المُحتوى أو مراقبته من داخل شركات التواصل عبر الإنترنت والعمليات الصحفية. وتحدد هذه التدابير من بين جملة من الأمور المُحتوى المسموح به على المنصة أو التطبيق، وإذا ما كان سيسمح له أن يبقى معروضاً إذا تم التحقُّق منه، والأهمية والبروز اللذين سيمنحان إياه وما إذا كان سيوضع أي إعلان بجانبه وإلى أي درجة ستتم التوصية به تلقائياً أو ما إذا كان سيتم الحد من نشره وما إذا كان سيخضع للتصنيف، وكذلك نوع المُحتوى المدفوع الذي سيظهر وبأي طريقة. وتتعلّق القضايا المطروحة هنا بالأحكام السياسية، وإنفاذها، ومسألة التصحيح. وهي ترتبط بكافة المحتويات على الإنترنت بما في ذلك المعلومات الصحيحة والمعلومات المضللة.

وتتم دراسة استجابات تنظيم المُحتوى من داخل شركات الإنترنت في المقام الأول من خلال النظر في سياسات هذه الشركات التي سوف يصار إلى تحليلها في هذا الفصل. وغالباً ما تؤدي هذه الاستجابات إلى اتخاذ تدابير تقنية أو خوارزمية سيتم تناولها بالتفصيل في القسم 6.2 أي الاستجابات التقنية/ الخوارزمية. كما وتشمل هذه الاستجابات عناصر معيارية وأخلاقية سيتم تناولها في القسم 7.1.

وتقوم المنظمات الإخبارية والصحفيون وغيرهم من الجهات الناشرة لمعلومات المصلحة العامة أيضاً باعتماد الاستجابات الخاصة بتنظيم المُحتوى بهدف التّصدي للتضليل. ويمكن أن تشمل هذه المهام تقديم التقارير المبنية على عمليات تعاونية في مجال التحقُّق من الوقائع، وتنظيم تحريري للمعرفة والموارد، وشراكات تعاونية للتحقُّق من الوقائع، وتنظيم المُحتوى في ما يخص المصادر والموارد، وتنظيم المُحتوى الذي ينتجه الجمهور (كالمُحتوى الذي ينتجه المستخدمون)، ومراقبة التعليقات. وتعالج الأقسام 4.1 و4.2 و5.3 جهود التنظيم التحريرية المرتبطة بالتحقُّق من الوقائع والتقارير الاستقصائية. وتتم معالجة القضايا الأخلاقية والمعيارية المرتبطة بالتنظيم التحريري للمُحتوى في القسم 7.1، كما ويتم تناول المبادرات التدريبية المتعلقة بتنظيم المُحتوى المضلل داخل المؤسسات الإعلامية في القسم 7.2 الذي يركز على الاستجابات التثقيفية.

في القسم أدناه، سوف يتم استعراض شروط الخدمة، والمبادئ التوجيهية المُجتمعية والسياسات التحريرية المتبعة في 11 شركة تواصل عبر الإنترنت (فيس بوك وانستغرام، وواتساب[181]، وغوغل، ويوتيوب[182] وتويتر، وويبو Weibo، وفيكاي VK، ووي تشات WeChat، ولاين وسناب تشات) لفهم كيفية تصدي هذه الشركات للتضليل بطريقة مباشرة أو غير مباشرة.[183] وتسعى هذه الأعمال إلى الحد من الجهات المتلاعبة، والسلوكيات المضللة، وما يعتبر مُحتوى ضاراً (François, 2019). ويتم التركيز على كيفية اتخاذ القرارات حول تنظيم المُحتوى/مراقبته، وإن كان يتم توظيف المستخدمين أو الأطراف الثالثة للمساعدة في مراقبة المُحتوى وكيف يتم ذلك، وما هي الآليات المتاحة لتقديم الطعون/التظلُّمات.

181 مُلاحظة: تطبيق واتساب مملوك من فيس بوك.

182 مُلاحظة: يوتيوب مملوك من غوغل.

183 مُلاحظة: انستغرام منصة مملوكة من فيس بوك.

وقد تكون الإجراءات المتخذة من قبل هذه الشركات فعالةً وديناميكيةً، ولكنها تطرح تساؤلات من قبل مختلف الجهات الفاعلة بشأن السلطات الممنوحة من خلال هذه العملية للجهات التجارية الخاصة. وقد برزت دعوات لتجاوز التنظيم الذاتي واعتماد آليات تنظيمية مستقلة وذلك بسبب مخاوف نشأت حول التطبيق العشوائي لتدابير التنظيم الذاتي – على سبيل المثال، تركيز الاستجابات على الولايات المتحدة والتخلي عن المسؤولية في بلدان الجنوب ذات المخاطر العالية (Ingram, 2018) – وتقليص التدابير لغايات الربح. وقد أدت أزمة المعلومات المضللة حول جائحة فيروس كورونا إلى تعزيز المخاوف من ألا تكون شركات التواصل عبر الإنترنت قادرةً على معالجة هذه المشكلة من خلال إجراءات تنظيم ذاتية أقوى (McNamee, 2020).

مثال آخر على ذلك هو السياسة المثيرة للجدل التي يتبعها فيس بوك والتي تعفي الإعلانات السياسية من التحقُّق من الوقائع (مراجعة القسمين 4.1 و7.1 اللذين يبحثان في سياسة فيس بوك المتعلقة بالتحقُّق من الوقائع والإعفاءات من ذلك). وقد تعرضت هذه السياسة للانتقاد على يد «الرئيس السابق لعمليات نزاهة الانتخابات» في الشركة في مقال بعنوان «عملت في مجال الإعلانات السياسية في فيس بوك» أو I worked on political ads at Facebook. وفي مقال عنوانه «يجنون الأرباح عبر التلاعب بنا» أو «They profit by manipulating us» نشرته صحيفة واشنطن بوست، كتب «يائيل أيزنستات» (Eisenstat 2019) «إنّ المشكلة الحقيقية تكمن في أنّ الأرباح التي يجنيها فيس بوك تنتج جزئياً عن طريق تضخيم الأكاذيب وبيع أدوات استهداف خطيرة تسمح للجهات السياسية الانخراط في حرب معلومات ذات مستوى جديد». وفي الآونة الأخيرة، سرّبت صحيفة وول ستريت جورنال تفاصيل من عرض لفريق فيس بوك حذّر فيه الفريق من مخاطر التنظيم الخوارزمي المتبع في الشركة. وقد ورد في العرض: «لدينا خوارزميات تستغل انجذاب الدماغ البشري إلى الانقسام... وإذا ما تركت هذه الخوارزميات من دون رادع [سوف يُعرض على المستخدمين] مزيد من المحتويات التي تُعزز الانقسام في محاولة لكسب اهتمامهم وزيادة الوقت الذي يمضونه على المنصة» (Horowitz & Seetharaman, 2020). وقد رد فيس بوك بالتالي: «إذا قامت الصفحات والمجموعات مراراً وتكراراً بمشاركة محتويات تنتهك معاييرنا المُجتمعية أو صنّفتها جهات التحقُّق من الوقائع على أنها كاذبة، نقوم بالحد من انتشارها، ونزيلها من التوصيات، كما ونحجب قُدرة الصفحات على جني الأرباح من المنشورات ووضع الإعلانات. كذلك، نقوم بإزالة صفحات ومجموعات بأكملها لنشرها محتويات مخالفة لقواعدنا بشكل متكرر». (Rosen, 2020).[184] من ناحية أُخرى وقبل قرار يوتيوب بإزالة حساب ناشر نظريات المؤامرة «أليكس جونز»، أكد موظف سابق في الشركة أن خوارزميات يوتيوب أوصَت بمقاطعه المعنونة «حروب المعلومات» أكثر من 15.000,000,000 مرة[185].

ولأن شركات التواصل عبر الإنترنت صُممت لتسمح بإيجاد مُحتوى معيّن و/أو الربط بين المستخدمين وتسهيل تنظيم ونشر المُحتوى والحث على التفاعل معه؛ من السهل استغلال ميزاتها لنشر التضليل. هذا يعني أن الأدوات التي صممت لتعزيز حرية التعبير والنفاذ إلى المعلومات يتم استخدامها كسلاح لمحاربة الحقيقة والصحة ولمنع النفاذ إلى معلومات موثوقة تتعلق بالمصلحة العامة (Posetti et al., 2019a). ومن الاستراتيجيات النموذجية التي يستخدمها مروجو التضليل لنشر المعلومات الخاطئة أو المضللة وضع عنوان جذاب لـ«قصّةٍ ما» لتجييش العواطف (والتي إما لا ترتقي إلى ما يُستشف من العنوان أو تكون مبنيةً

[184] ومع ذلك عند كتابة هذا التقرير؛ كان فيس بوك مستمراً في عدم إخضاع الإعلان السياسي والمُحتوى المنشور من قبل السياسيين لتدابير مكافحة التضليل هذه، كما ذكَر هذا التقرير سابقاً.

[185] https://twitter.com/gchaslot/status/967585220001058816s=21

على معلومات ملفقة) وذلك لحث المستخدمين على التفاعل والنقر. وهذا ما يُعرف بالعناوين الخاطفة أو clickbait. وقد ثبت أن المُحتوى المثير للعواطف عادةً ما يميل إلى خلق تفاعل أكبر (Martens et al., 2018). وعبر اجتذاب المستخدمين ليتفاعلوا ويعجبوا ويشاركوا المحتويات المضللة يمكن للجهات المضللة أن تستغل التشبيك الذي توفره خوارزميات المنصات المصممة لإبراز محتويات تهم المستخدمين وبالتالي لتسريع وتوسيع نطاق انتشار الرسائل الكاذبة (DiResta, 2018). ويناقش القسم 6.2 التدابير التي تعتمدها شركات التواصل عبر الإنترنت للحد من انتشار العناوين الخاطفة.

وفي اقتصاد الإعلان الرقمي، تلعب هذه الشركات، بفعل الأمر الواقع، دور «وسطاء جذب انتباه» (Wu, 2017) وعليها أن تقيم توازناً صعباً؛ علماً أنّ المُحتوى الذي يتمتع بأكبر قدر من التفاعل هو أيضاً الأكثر ربحاً في مجال جمع البيانات و/أو وضع الإعلانات ذات الصلة. والجدير بالذكر أنّ خروقات البيانات (على سبيل المثال كما في حال كامبريدج أناليتيكا)، والتدخل الأجنبي في الانتخابات الديمقراطية كما في الولايات المتحدة في عام 2016 (Howard et al., 2018)، ونشر المعلومات المضللة بشكل كبير عبر تطبيقات المراسلة خلال الانتخابات (كما في الهند في عام 2019) والأزمات الصحية على غِرار جائحة فيروس كورونا (مُراجعة النقاش حول ذلك أدناه)، قد وضعت ضغوطاً على الشركات دفعتها إلى اتخاذ إجراءات للحد من انتشار المُحتوى المضلل على خدماتها (Burgos, 2019).

ويعتمد الكثير من التواصل على الإنترنت على الوسطاء، ولذلك يخضع لسياساتهم. ويمكن أن يشمل هؤلاء أشخاصاً يقومون بالتحقُّق من الوقائع، ومراقبين ومحققين، بما في ذلك أشخاص توظفهم الوكالات الإخبارية وشركات التواصل عبر الإنترنت، والشراكات القائمة بين المنظمات الإخبارية وخبراء التحقُّق الآخرين[186]. كذلك تخضع عمليات التواصل هذه للبنية الرقمية للشركة؛ أي للبروتوكولات التقنية التي تمكّن وتقيّد وتحدد سلوك المستخدمين على الإنترنت، والتي تعكس نماذج الأعمال واعتبارات أخرى. وتختلف هذه الميزات التقنية والتصميمية من خدمة إلى أخرى. وبشكل ملموس، ما يحدد ملامح التواصل (والبحث والاكتشاف) على منصة معيّنة هو كيف يتم إنشاء الروابط بين الحسابات على وسائل التواصل الاجتماعي وكيف تتم المحافظة عليها وكيف يتم التفاعل بين المستخدمين من خلال التكنولوجيا، إضافةً إلى استخدام مصفاة المُحتوى والتحويل الخوارزمي إلى بيانات. (Bossetta, 2018). وهذه الاختلافات في الغرض والبنية الهندسية تفسّر أيضاً بشكل جزئي اختلاف استراتيجيات تنظيم المُحتوى من شركة إلى أخرى.[187] وسيتم البحث في هذه التوترات والتحديات التي تواجه استخدام استجابات تنظيم المُحتوى لتعزيز حرية التعبير والدفاع عنها بمزيد من التفصيل في التقييم الوارد في القسم الأخير من هذا الفصل.

والمفتاح لاستجابات ناجحة في مجال تنظيم المُحتوى هو وجود رقابة مستقلة. وفي هذا السياق، لمنظمات المُجتمع المدني والمواطنين دور مهم، حيث أن بإمكانهم التحقُّق من الطرق التي تقوم من خلالها منصات التواصل الاجتماعي بحماية حرية التعبير وتطبيق الشفافية الكاملة في استجاباتها

[186] مراجعة القسمين 4.1 (الرصد والتحقُّق من الوقائع) و5.3 (الاستجابات الانتخابية) لنقاش مفصل حول الدور التنظيمي لعملية التحقُّق من الوقائع.

[187] على سبيل المثال، خلال الاستفتاء الآيرلندي في عام 2018 على التعديل السادس والثلاثين من القانون الدُّستوري (حول الإجهاض)، قرر غوغل عدم نشر الإعلانات السياسية، في حين قام فيس بوك بحظر إعلانات الجهات الخارجية فحسب. وبناءً على سياسته الدعائية، حظر تويتر الإعلانات حول الإجهاض منذ البداية (O'Brien & Kelly, 2018; Satariano, 2018)

الساعية إلى تنظيم المُحتوى. وتُعتبر الشفافية، والمُساءلة، وإمكانية الطعن في ما يخص إجراءات تنظيم المُحتوى أموراً ضرورية لحماية حرية التعبير خاصةً وأن هناك إمكانية لارتكاب الأخطاء؛ إن كان من خلال خوارزميات المنصات أو من خلال المراقبين. ونظراً لأعداد المستخدمين والمنشورات اليومية الهائلة، قد تُشكّل آثار تنظيم المُحتوى مشكلةً كبيرةً إذا لم تُكبح وبقيت على حالها من دون رادع.

1.1.6 مقاربة شركات التواصل عبر الإنترنت لتنظيم المُحتوى

يقدم هذا القسم نظرةً عامةً حول الطريقة التي تتبعها شركات التواصل عبر الإنترنت لتنظيم أو مراقبة المُحتوى والحسابات بناءً على شروط الخدمة والمبادئ التوجيهية المُجتمعية والسياسات التحريرية.[188]
وفي ما يلي قائمة بالمصادر الأولية المستخدمة في تحليل كل منصة:

فيس بوك وانستغرام:

https://www.facebook.com/communitystandards/introduction; https://help.instagram.com/477434105621119; https://transparency.facebook.com/; Facebook & Instagram (2019)

واتساب:

https://www.whatsapp.com/legal/?eea=0#terms-of-service; https://faq.whatsapp.com/21197244/#Report; https://blog.whatsapp.com/10000647/More-changes-toforwarding
WhatsApp (2019)

غوغل ويوتيوب:

https://about.google/community-guidelines/; https://transparencyreport.google.com;
https://www.youtube.com/yt/about/policies/#community-guidelines;
Google and YouTube (2019)

تويتر:

https://help.twitter.com/en/rules-and-policies#research-and-experiments;
https://transparency.twitter.com/en/removal-requests.html;
Twitter (2018); Twitter (2019)

VK:

https://vk.com/licence ; https://vk.com/blog ; https://vk.com/support?act=home;
https://vk.com/help?page=cc_terms ; https://vk.com/page-76477496_50995734

Weibo:

https://www.weibo.com/signup/v5/protocol

[188] نتوجه بالشُّكر إلى زملائنا الباحثين الذين أعطونا من وقتهم الثمين لقراءة وتحليل شروط الخدمة والمبادئ التوجيهية المُجتمعية والسياسات التحريرية لمختلف التطبيقات: Weibo وWeChat الزميلة Olivia Sie وتطبيق VK الزميل Vsevolod Samokhvalov وتطبيق LINE الزميل Koji Yamauchi في اللُّغات الأكثر استخداماً على المنصات؛ أي الصينية والروسية واليابانية.

:WeChat

https://www.wechat.com/en/service_terms.html; (للمستخدمين الأجانب)
https://weixin.qq.com/cgi-bin/readtemplate?lang=en&t=weixin_agreement&s=default&cc=CN
(للمستخدمين في الصين)؛
https://help.wechat.com/;
https://wechatwiki.com/wechatresources/wechat-rules-and-marketing-restrictions/

LINE:

https://terms.line.me/line_terms/?lang=ja; LINE (2020)

سناب تشات:

https://www.snap.com/en-US/community-guidelines ; https://www.snap.com/en-US/adpolicies/political/; https://businesshelp.snapchat.com/en-US/article/political-ads-library

ينصبّ التركيز على شركات التواصل عبر الإنترنت (وسائل التواصل الاجتماعي، وتطبيقات المراسلة، وتطبيقات مشاركة مقاطع الفيديو، ومُحركات البحث)، حيث كانت في صميم المطالبة بالتّصدي للتضليل على الإنترنت. وفي مراجعة شروط الخدمة، والمبادئ التوجيهية المُجتمعية والسياسات التحريرية («قواعد المنصة»)، يمكن ملاحظة وجود الاستجابات التالية في مجال تنظيم المُحتوى:

1. التنبيه من مُحتوى معيّن ومراجعته.
2. استخدام مصفاة مُحتوى معيّن والحد منه أو حجبه أو إزالته.
3. الترويج للمُحتوى وخفض ترتيبه.
4. تعطيل أو إزالة حسابات.
5. اعتماد الشفافية في المحتويات المدعومة.
6. إشراك المستخدمين.
7. آليات الطعن/التظلُّم.

في الجدول أدناه، تم مسح إجراءات اتخذتها 11 شركةً عالميةً موزعةً جغرافياً تتمتع بقاعدة كبيرة من المستخدمين. في تحاليل لاحقة، تم تفصيل الاختلافات في تنظيم المُحتوى والحسابات بين هذه الشركات، كما وتم إعطاء أمثلة. ويستند التحليل إلى وثائق (سياسات، تدوينات، تقارير شفافية) متعلقة بتنظيم المُحتوى قدمتها شركات التواصل عبر الإنترنت. وفي الجدول تم إدراج الإجراءات التي عُثر على أدلة حولها في الوثائق. ومتى لم يتم العثور على أيّ دليل حول هذا الإجراء، وفي حال عدم كفاية الأدلة؛ تُركت خانة الإجراء فارغةً. ومتى تم وضع الإجراء بين قوسين، يعني أنه يعتمد على نوع المُحتوى أو المستخدم.

تنظيم المُحتوى/ الحساب		فيس بوك وانستغرام	واتساب	حملات غوغل ويوتيوب	تويتر	VK	Weibo	WeChat	LINE	سناب تشات
تنبيه ومراجعة المُحتوى	آلياً	x		x	x	x	x	x	x	
	بشرياً	x		x	x	x	x	x	x	x
	مراجعة من قبل طرف ثالث	x	(x)	x	x	x				
	مُستشار خارجي	x		x						
استخدام مصفاة المحتوى الحد من المُحتوى، أو حجبه أو إزالته	تحميل المصفاة	x		x	x	x	x	x		
	إعادة إرسال المُحتوى المحظور		x							
	قيود على أساس:									
	قواعد الشركة	x	x	x	x	x	x	x	x	(x)
	إنفاذ القانون	x	x	x	x	x	x	x	x	x
ترويج أو خفض مرتبة المُحتوى	الترويج للمصادر الموثوقة	x		x	x	x				(x)
	تخفيض مرتبة العناوين الخاطفة clickbait أو المُحتوى غير المتفق عليه	x		x	x	x		x		
تعطيل أو تعليق الحسابات	مقاربة متدرجة:									
	تحذير			x						
	ميزات محدودة	x	x	x	x		x	x	x	
	تعليق	x	x	x	x	x	x	x	x	x
شفافية في المُحتوى	تحديد المُحتوى المدعوم	x		x	x	x	x	x	x	x
	مركز الشفافية الإعلانية	x		x	x					x
إشراك المستخدمين	يمكن للمستخدمين التنبيه من مُحتوى لتتم مراجعته	x	x	x	x	x	x	x	x	
	يمكن للمستخدمين أن يحجبوا مُحتوى/ حساب معيّن أو تعليقه لفترة معيّنة	x	x	x	x	x	x	x	x	x
	يمكن للمستخدمين تحديد أولوية المُحتوى/الحسابات	x		x	x	x	x	x	x	x
	يمكن للمستخدمين تغيير الفئات التي عيّنوها للإعلان	x		x	x				x	
يمكن للمُستخدمين تغيير الفئات التي عيّنوها للإعلان	إشعار باتخاذ إجراءات	x		x	x			(x)	x	
	إمكانية الطعن/التظلُّم	(x)		(x)	(x)	(x)			x	
	التبليغ عن القرار بشأن الطعن/التظلُّم	(x)		(x)	(x)	(x)				

الجدول 5. استجابات تهدف إلى تنظيم المُحتوى من قبل شركات التواصل عبر الإنترنت

ماذا تتضمن شروط الخدمة والمبادئ التوجيهية المُجتمعية والسياسات التحريرية التي تتبعها هذه الشركات؟ هي تعطي تفاصيل حول أي نوع من المُحتوى يؤدي إلى اتخاذ إجراءات – بدءً من السُّلوك العنيف والإجرامي (العنف والتحريض عليه والأفراد والمنظمات التي تُعتبر إرهابيةً أو إجراميةً أو الترويج للجريمة أو الدعاية لها/تنسيق الضرر، وانتهاكات السلع الخاضعة للتنظيم، والغش والخداع، والتدخل في الانتخابات) إلى المُحتوى المعترض عليه (خطاب الكراهية والعنف والمُحتوى القاسي، العري والنشاط الجنسي للبالغين، والاستدراج الجنسي والمعاملة القاسية والسيئة، والتنمُّر)، وأكثر من ذلك.

1. التنبيه أو التحذير من مُحتوى معيّن ومُراجعته

يمكن التنبيه من المحتويات المسيئة أو غير القانونية التي تُنشر على منصات التواصل على الإنترنت من خلال الوسائل الآلية، أو يدوياً من قبل المستخدمين أو من قبل أطراف ثالثة (مثل سُلطات إنفاذ القانون وجهات التحقُّق من الوقائع، والوكالات الإخبارية العاملة في إطار شراكة). ويبدو أن الكشف الآلي آخذ في الارتفاع؛ وهو مهم للتصدّي للجهود الهادفة إلى نشر المعلومات المُضلّلة وأنواع أُخرى من المحتويات الضارة (مراجعة القسم التالي 6.2 الاستجابات التقنية/الخوارزمية). ولتوضيح آليّة تنظيم المُحتوى، تم خلال الفترة الممتدة من تموز/يوليو إلى أيلول/سبتمبر 2019، إزالة نحو 8765893 مقطع فيديو من يوتيوب. ومن بين هذه المقاطع، كان قد تم الإبلاغ عن 602826 مقطع فقط من قبل أفراد[189]. بالنسبة إلى الكشف البشري، يفتخر تويتر، على سبيل المثال، بتوفير بوابة لدعم الشركاء Partner Support Portal، وهي خط سريع يسمح لشركاء تويتر بتلقي الردود على التقارير وتسهيل تبادل المعلومات (Twitter, 2019). وللمنصات الأخرى شركاء متميزون أيضاً تتعاون معهم خاصةً سُلطات إنفاذ القانون.

وتوظّف معظم المنصات الإلكترونية أشخاصاً يقومون بمراجعة المُحتوى. وقد زاد كل من فيس بوك وغوغل على وجه الخصوص من عدد الموظفين العاملين في مجال مراقبة المُحتوى على مر السنين. وقد استخدم فيس بوك 15000 موظف لمراجعة المُحتوى في 2019 (Facebook, 2019)، في حين أعلن غوغل في عام 2017 أنه سيقوم بتوظيف 10000 مشرف على المُحتوى بحلول نهاية عام 2018 (Hoggins, 2019). وفي عام 2019 كان طاقم مراقبة المُحتوى في تويتر يتألّف من حوالي 1500 موظف. ويعمل معظم مراقبو المُحتوى على المنصات الإلكترونية كمتعاقدين خارجيين (Dwoskin, Whalen & Cabato, 2019). وفي VK، يعمل فريق مؤلّف من أكثر من 100 شخص، مُقسّم إلى عدة مجموعات بناءً على خصائص ‹الانتهاكات› التي تم كشفها في عمليات تنظيم المُحتوى.

كما ويُستخدم التعلُّم الآلي أيضاً لكشف المعلومات المضللة والبريد الإلكتروني غير المرغوب به.[190] ومع ظهور جائحة فيروس كورونا انتقلت معظم هذه الشركات نحو الاستخدام المكثف للتنظيم الآلي للمُحتوى. وتناقش مسألة «التشغيل الآلي» لتنظيم المُحتوى في القسم التالي (6.2) الذي يركّز على الاستجابات التقنية والخوارزمية.

وفي بلدان أُخرى، أنشأت منظمات التحقُّق من الوقائع حسابات خاصةً للتنبيه من المعلومات المشتبه بزيفها ليُصار إلى التحقُّق منها. وعلى الرغم من أنّ بعض هذه الخدمات قد يكون مدعوماً من قبل

....

[189] https://transparencyreport.google.com/youtube-policy/removals

[190] https://vk.com/page-76477496_50995734

شـركات التواصـل عبـر الإنترنـت، لا تقـوم هـذه الشـركات بـإدارة هـذه الخدمـات بشـكل مباشـر وهـي تُعالـج محتويـات أوسـع مـن تلـك التـي تحيلهـا إليهـا الشـركات (Tardáguila, 2019).

وأخيـراً، يعمـل فيـس بـوك وغوغـل مـع جهـات خارجيـة علـى غِـرار المستشـارين القانونييـن، عنـد الضـرورة، للتحقُّـق ممـا إذا كان مُحتوى معيّـن ينتهـك المعاييـر و/أو التشـريعات (خاصـةً الوطنيـة منهـا). وفـي عـام 2019، أعلـن فيـس بـوك عـن خطـة لإنشـاء مجلـس رقابـة يهـدف إلـى «حمايـة حريـة التعبيـر مـن خـلال اتخـاذ قـرارات مسـتقلة وذات مبـدأ حـول محتويـات مهمـة وعبـر إصـدار نصائـح سياسـية حـول سياسـات المُحتـوى الـذي يتبعهـا فيـس بـوك» (Facebook, 2019). وقـد تـم الإعـلان عـن أول أعضـاء فـي المجلـس فـي أيار/مايـو 2020 (Facebook, 2019). وتـم الإشـادة بالمجلـس فـي بعـض الأوسـاط علـى أنـه «المحكمـة العليـا للمُحتـوى»، حيـث كان مـن المتوقـع أن يقـوم بالحـد مـن سياسـة فيـس بـوك التـي تسـمح بالأكاذيـب فـي الإعلانـات السياسـية (Levy, 2020). ولكـن فـي تموز/يونيـو 2020، تبيّـن أنّ اختصـاص المجلـس يقتصـر علـى مراجعـة القـرارات بإزالـة المُحتـوى (مراجعـة المناقشـة فـي القسـم 7.1 حـول الاسـتجابات المعياريـة والأخلاقيـة للمعلومـات المضللـة). مـن جانبـه، أنشـأ تويتـر مجلـس الثقـة والسـلامة[191] الـذي يقـدم المشـورة حـول تطبيـق قواعـد السـلامة الـذي يتبعهـا تويتـر.

2. استخدام مصفاة المُحتوى وإزالته وحجبه وغيرها من القيود المفروضة على المُحتوى:

يتـم تطبيـق تدخـلات تؤثـر علـى توافـر المُحتـوى بنـاءً علـى أسـاس شـروط خدمـة الشـركات، أو مبادئهـا التوجيهيـة المُجتمعيـة وسياسـاتها التحريريـة أو المتعلقـة بإنفـاذ القانـون (مراجعـة القسـم 5.1 حـول الاسـتجابات التشـريعية ومـا قبـل التشـريعية والسياسـتية). ويمكـن أن نلاحـظ أن هـذه القواعـد قـد تكـون أكثـر تقييـداً مـن أساسـها القانونـي فـي عـدد مـن الـدول. وخيـر مثـال هـو قـرار تويتـر حظـر الإعلانـات السياسـية المدفوعـة علـى الصعيـد العالمـي فـي تشـرين الثانـي/نوفمبـر 2019. فـي المقابـل، قـرر فيـس بـوك مواصلـة نشـر الإعلانـات السياسـية (مُراجعـة القسـم 4.1 أعـلاه) مـن دون التحقُّـق مـن صحـة مضمونهـا، وقـاوم الدعـوات لمنـع اعتمـاد الاسـتهداف الدقيـق فـي هـذه الإعلانـات. وظهـر اختـلاف المقاربـات بيـن الشـركتين جليـاً فـي خـلاف علنـي اندلـع بيـن تويتـر وفيـس بـوك والرئيـس الأمريكـي السّـابق دونالـد ترامـب فـي أيار/مايـو وحُزيران/يونيـو 2020 بعـد أن وضـع تويتـر تنبيهـاً علـى تغريـدة للرئيـس السّـابق واصفـاً إياهـا بالمُضلّلة، وكانـت حـول بروتوكـولات الانتخابـات (Hatmaker, 2020) كمـا وقـام بإخفـاء إحـدى تغريداتـه «لتمجيدهـا للعنـف» (BBC, 2020c). وأوضـح الرئيـس التنفيـذي لشـركة فيـس بـوك «مـارك زوكربيـرغ» أنّ فيـس بـوك لـن يتخـذ إجـراءً مماثـلاً ضـد شـخصية سياسـية مهمـة لأن الشـركة ترفـض أن تكـون «حكمـاً فـي الحقيقـة». (للاطـلاع علـى مناقشـة أكثـر تفصيـلاً لهـذه المسـألة، مراجعـة القسـم 7.1 حـول الاسـتجابات المعياريـة والأخلاقيـة).

وتتميـز المصفـاة بأنهـا تتـم اسـتباقياً أي قبـل نشـر وتوزيـع المُحتـوى. أمـا القيـود، مـن منـع وإزالـة منشـورات معينـة وتوزيـع المُحتـوى، فقـد تتـم أيضـاً بأثـر رجعـي، أي بعـد أن يتـم نشـر المُحتـوى. وفـي مـا يتعلـق بتفعيـل مصفـاة المحتـوى (قبـل النشـر)، تسـتخدم المنصـات قواعـد بيانـات البعثـرة[192] المـزودة «ببصمـات رقميـة» للمحتويـات التـي تـم التنبيـه منهـا، وللمحتويـات الإرهابيـة، وتلـك التـي تحتـوي علـى صـور تمثـل إسـاءةً جنسـيةً

---

[191] https://about.twitter.com/en_us/safety/safety-partners.html

[192] تُشير «بعثرة قواعد البيانات» أو ما يُعرف بـ Hashing إلى استخدام «قيمة بعثرة» أو مرجعاً لتصنيف واسترجاع وحماية الأمور الموجودة في قاعدة بيانات معيّنة (Zilavy, 2018).

ضد الأطفال، وللمُحتوى الذي ينتهك حقوق المؤلف؛ وذلك ليُصار إلى كشف هذه المحتويات ومنع إعادة تحميلها. وفي هذا السياق، أنشأ كل من يوتيوب وفيس بوك ومايكروسوفت، وتويتر منتدى الإنترنت العالمي لمكافحة الإرهاب[193] (GIFCT) للتعاون في وضع حلول تكنولوجية لمكافحة التطرف العنيف على منصات هذه الشركات.[194]

وتقوم شركات التواصل عبر الإنترنت أيضاً بإزالة وحجب أو الحد من مُحتوى معيّن عند تلقّيها إخطارات آلية أو بشرية حول مواد تم الاعتراض عليها، فتطبق تدابير يختلف مقياسها باختلاف طبيعة الانتهاك. ويُعتبر الحد من انتشار المعلومات المضللة على واتساب صعباً بشكل خاص، نظراً لتشفير المحادثات. وقد بدأ واتساب بالحد من عدد المرات التي يمكن مشاركة الرسائل إلى خمس مرات. وقد اعتُمدت هذه الميزة للمرة الأولى في الهند في تموز/يوليو من عام 2018، ثم تم نشرها في جميع أنحاء العالم في كانون الثاني/يناير 2019 ,WhatsApp) (2019a. وتم تشديد القيود المفروضة بكثافة على إعادة الإرسال خلال جائحة فيروس كورونا، حيث قلّص واتساب عدد المرات التي يمكن إعادة ارسال رسالة متكررة إلى مرة واحدة فقط (El Khoury, 2020). (مراجعة المناقشة أدناه حول «العواقب غير المقصودة» لهذه القيود). وليس من الواضح ما إذا كان تبادل البيانات الوصفية بين واتساب والشركة الأُم فيس بوك سيساعد أيّ من الشركتين على مكافحة التضليل.

أما الشركة الصينية WeChat فتعمل وفقاً لاتفاقيتين للخدمات -الأولى للصين والأخرى للمستخدمين الدوليين. وقد أظهر بحث طولي أجراه مُختبر المواطنين في «جامعة تورونتو» أنّ WeChat يقوم بمراقبة المُحتوى في الوقت الحقيقي، فيقوم بإزالة المُحتوى على أساس سلاسل مكونة من كلمات أساسية وعناوين إلكترونية وصور. ووجد الباحثون أيضاً أنه تتم استخدام مصفاة الرسائل من المستخدمين الصينيين بوتيرة أعلى من مصفاة رسائل المستخدمين الدوليين. كذلك، يتم تفعيل مصفاة المُحتوى المنشور عبر لحظات WeChat ومجموعات المُحادثة فيه بوتيرة أعلى مقارنةً بمصفاة المحادثات الفردية (Ruan et al., 2016; Knockel & Xiong, 2019).

3. الترويج للمُحتوى وخفض ترتيبه

يستند خيار آخر اعتمدته شركات التواصل عبر الإنترنت إلى افتراض مفاده أن «حرية التعبير ليست حرية الوصول» (DiResta, 2018). فيتم بناءً على ذلك الترويج للمصادر الموثوقة وفقاً لمعايير معينة عبر الخوارزميات فيما يتم خفض مرتبة المُحتوى المُضلّل (أو ينشر الكراهية أو يحتمل أن يكون ضاراً بطرق أخرى) على شريط الأحداث الجارية. (مراجعة القسم 7.3)

وعلى فيس بوك، يتم التعامل مع العناوين الخاطفة أو clickbait عن طريق الحد من ظهور المُحتوى متى كان يحمل عنواناً[195] «يحجب معلومات أو يبالغ فيها» (Babu, Lui & Zang, 2017). وقد التزم فيس بوك أيضاً بالحد من ظهور المقالات التي تم التحقُّق من صحتها على يد منظمات شريكة، وتبيّن أنها تُعاني من شوائب. وتضيف الشركة إطاراً على هذه المقالات من خلال وضع مقالات تم التحقّق من صحتها تحت المقالات المضللة

193 https://www.gifct.org

194 https://transparencyreport.google.com/youtube-policy/featured-policies/violent-extremism?hl=en

195 https://about.fb.com/news/2017/05/news-feed-fyi-new-updates-to-reduce-clickbait-headlines/

المعينة[196]. (لكن، وكما ذُكر في الأقسام 4.1، 7.1، و5.3، لا تخضع فئات معينة من الإعلانات السياسية لجهود التحقُّق من الوقائع هذه). بالإضافة إلى ذلك، بدأت الشركة بالدفع لمجموعة مختارة من الوسائل الإخبارية لتقديم مُحتوى يتم عرضه في قسم منفصل مخصص «للأخبار». وعند كتابة هذا التقرير، كان لا يزال يتم اختبار هذا الموضوع (أي في نسخة بيتا) ولم يكن متاحاً إلا لنحو مئات آلاف المستخدمين في الولايات المتحدة (Kafka, 2020). أما يوتيوب فيعطي الأولوية للمُحتوى الصادر عن وكالات إخبارية موثوقة ويضعه في خانات «أهم الأخبار» و«الأخبار العاجلة» كإجراء تنظيمي يهدف إلى تسليط الضوء على المُحتوى الموثوق به، على الرغم من أنّ هذا متاح حالياً فقط للمستخدمين في الولايات المتحدة (Google & YouTube, 2019).

ويميّز سناب تشات نفسه عن منصات التواصل الاجتماعي الأخرى من خلال «الفصل بين المواد الاجتماعية والمواد الإعلامية» (Spiegel, 2017). فتخصص المنصة قسماً منفصلاً «هو اكتشاف سناب تشات»، يعمل على عرض قصص من جهات إخبارية ومنتجي محتويات وجهات مُجتمعية مختارة بشكل خوارزمي، ويخضع للمراقبة والترويج من قبل المحررين في سناب تشات (Snapchat, 2017). وفي عام 2020، ألغت الشركة شريط الأحداث الجارية التابع لرئيس الولايات المتحدة من قسم الاكتشاف (Newton, 2020).

4. تعطيل وتعليق الحسابات

بالإضافة إلى تنظيم المُحتوى، تعمد شركات التواصل عبر الإنترنت إلى التّصدي لما تسميه سلوكاً ومُحتوى غير أصيل (مُفتعل) أي زائفاً على مستوى حسابات المستخدمين. ويمكن نشر التضليل على الإنترنت بسهولة من خلال حسابات يتم اختراقها أو إعدادها خصّيصاً لمثل هذه السُّلوكيات، وغالباً بأعداد وفيرة، لغرض التلاعب بالمحتويات. ويحظر العديد من الشركات «السلوك المُفتعل المعد له» (بما في ذلك التدخل من قبل حكومات أجنبية) في شروط الخدمة. ويشير فيس بوك إلى أنّ معالجة هذه السلوكيات هو تحد مستمر يلتزم بالتّصدي له من خلال العمل «بشكل مستمر على تحسين أدائه في هذا المجال من خلال بناء تكنولوجيات أفضل، وتوظيف المزيد من الأفراد، والعمل بشكل وثيق مع سلطات إنفاذ القانون وخبراء الأمن والشركات الأخرى» (Facebook and Instagram, 2019). في ضوء ذلك، عمدت الشركة إلى تحديث سياساتها المتعلقة بالسلوك غير الأصيل/المُفتعل (CIB) في تشرين الأول/أكتوبر من عام 2019، موضحةً كيف تعمل ضد «مجموعة من الأنشطة المُفتعلة أجنبيةً كانت أم محلية، حكوميةً أو غير حكومية» (Gleicher, 2019).

وتتدخل بعض الشركات خلال عملية تسجيل الحساب وبعد ذلك طول مدة حياته. فعلى سبيل المثال، قام واتساب في فترة تمتد إلى ثلاثة أشهر «بحظر أكثر من مليون حساب في الشهر بسبب سلوكيات آلية (غير بشرية) بأعداد وفيرة». كما وتم حظر 20% من هذه الحسابات في مرحلة التسجيل (WhatsApp, 2019a). ويميل تنظيم المُحتوى على منصة معيّنة إلى اتباع مقاربة متدرجة تقضي بتوجيه التحذيرات قبل فرض عقوبات. فتقوم «شركة لاين» Line [197] والعديد من الشركات الأخرى، باستثناء VK [198] وسناب تشات[199] بتعطيل حساب المستخدم بشكل مؤقت وبعدها تقوم بوقفه، إذا تبيّن أنه يقوم بانتهاك بنود وشروط الاستخدام و/أو القوانين.

196 https://www.facebook.com/help/1952307158131536

197 https://terms.line.me/line_terms/?lang=ja

198 https://vk.com/licence

199 https://www.snap.com/en-US/community-guidelines/

كذلك، يقوم فيس بوك بتعليق صفحات مجموعات معيّنة تنتهك شروط الخدمة، سواءً على موقعه أو على انستغرام. ومن الأمثلة الحديثة على ذلك إزالة 790 مجموعةً على فيس بوك تابعةً لحركة QAnon و100 صفحة و1500 إعلان. كما وتم تقييد 1950 مجموعةً أخرى على انستغرام (Facebook, 2020). واعتبرت مصادر نظرية المؤامرة هذه على أنها تنتهك سياسات فيس بوك لأنها تحتفي بأعمال العنف و»لها متابعين لهم أنماط من السلوك العنيف». وشمل ذلك أيضاً 440 صفحةً وأكثر من 10000 حساب على انستغرام مرتبط بحركة QAnon المتطرفة (Facebook, 2020b). وأتى هذا بعد تحقيق داخلي قامت به الشركة أظهر أن عضوية هذه المجموعات تتجاوز 3 ملايين شخص (Sen & Zadrozny, 2020).[200]

وليس من تمييز واضح بين تعليق الحسابات على أساس قيامها بسلوك زائف وتعليقها لتبادل المُحتوى المضلل، حيث أنّ المفهومين يتداخلان في كثير من الأحيان في المبادئ التوجيهية المُجتمعية. ويطبق يوتيوب سياسات أوسع نطاقاً في هذا المجال. فإذا تم العثور على انتهاكات للمبادئ التوجيهية المُجتمعية، يقوم يوتيوب بإزالة المُحتوى وتوجيه تحذير للحساب. وبعد 3 تحذيرات خلال فترة 90 يوماً يتم وقف الحساب. وتوضع التحذيرات على قناة يوتيوب المستهدفة، ولكن من دون عواقب إضافية. ويترتب عن هذه التحذيرات تعطيل قدرة أصحاب الحسابات على تحميل وإنشاء وتحرير المُحتوى على موقع يوتيوب لمدة أسبوع واحد (التحذير الأول)، ثم لأسبوعين اثنين (التحذير الثاني) مما يؤدي في النهاية إلى حذف قناة يوتيوب (بعد التحذير الثالث). ومع ذلك، في الحالات التي تتطلب التدخل في انتهاكات هي أبعد من انتهاكات المبادئ التوجيهية المُجتمعية (على سبيل المثال رداً على «شكوى من أحدهم حول انتهاك الخصوصية أو بناءً على أمر من المحكمة»)، لا تنطبق سياسة التحذير هذه بل قد يصار إلى تعليق الحساب بشكل فوري.[201]

5. الشفافية في تنظيم المُحتوى والمُحتوى المدعوم

بما أنّ مواقع وتطبيقات التواصل الاجتماعي أصبحت تعتبر بشكل متزايد من الحيّز العام على شبكة الإنترنت، يؤكد البعض على أنّ تنظيم المُحتوى قد يؤثر على الحق في حرية التعبير. وعلى الرغم من أن الجهات الخاصة تتمتع بالحق في اتخاذ القرارات التي تتعلّق بسياسات تنظيم المُحتوى على خدمتها (ضمن الحدود القانونية)، يبقى حق الفرد قائماً في الخضوع لإجراءات قانونية مُحقّة. إضافةً إلى ذلك، لابد من إعطاء معلومات معيّنة واعتماد الشفافية في إعلام المستخدمين والأطراف الثالثة بالطرق التي تُتخذ فيها القرارات، وذلك لضمان أن تؤخذ هذه القرارات بناءً على أُسُس قانونية و/أو عادلة. وفي عام 2018، قامت مجموعة في الولايات المتحدة معنيّة بحرية التعبير في مجال تنظيم المُحتوى على الإنترنت ومكوّنة من أكاديميين ومدافعين عن الحقوق في المجال الرقمي بوضع مبادئ سانتا كلارا حول الشفافية والمُساءلة في تنظيم المُحتوى.[202] وتضع هذه المبادئ معايير عالية؛ على الشركات التقيُّد بها، وتقترح معايير مفصّلة لتقارير الشفافية والتحذيرات والطعون. في الواقع، وكحيّز عام، لابد للشركات المُهيمنة أن تطبق معايير دولية لا معايير محصورة وأكثر محدودية.

[200] مُراجعة النقاش الذي استضافته منظمة Columbia Journalism Review حول نظرية المؤامرة والتضليل اللذين تشرهما حركة QAnon، والذي ضم مجموعةً من الصحفيين والباحثين: https://galley.cjr.org/public/conversations/-MFpKx9fqAg-5dUs2DirW

[201] https://support.google.com/youtube/answer/2802032?hl=en

[202] https://santaclaraprinciples.org/

يقوم كل من فيس بوك/انستغرام[203] وغوغل ويوتيوب[204] وتويتر[205] وسناب تشات[206] ولاين[207] بتقديم تقارير شفافية دورية (على سبيل المثال بشكل فصلي) حول مُمارسات تنظيم المُحتوى وتماشيها مع المُتطلّبات الخارجية (القانونية). وتميل هذه التقارير إلى أن تكون أقل شفافيةً في ما يخص العمليات والمُمارسات الداخلية. كل هذه الشركات ما عدا لاين LINE تحتفظ بمكتبات للإعلانات (السياسية). وتغطي مكتبات فيس بوك وتويتر جميع الإعلانات عالمياً، في حين يوفر غوغل تقارير عن الإعلانات السياسية في الاتحاد الأوروبي والهند والولايات المتحدة، ويغطي سناب تشات الإعلانات السياسية في الولايات المتحدة. وتجدر الإشارة إلى أنّ الأرجنتين وكندا والاتحاد الأوروبي وفرنسا والهند تُلزم شركات الخدمات عبر الإنترنت (والمرشحين للانتخابات) باعتماد الشفافية في الإعلانات السياسية. وتعد هذه استجابةً سياستيةً ترددها دول عديدة، بما في ذلك أستراليا وبلجيكا وآيرلندا وإسرائيل وإيطاليا وهولندا ونيوزيلندا ومنظمة الدول الأمريكية والمملكة المتحدة والولايات المتحدة (مراجعة القسم 5.1 حول الاستجابات التشريعية وما قبل التشريعية والسياستية).

ولكن منذ 22 تشرين الثاني/نوفمبر 2019، يقوم تويتر بحظر الإعلانات السياسية عالمياً والإعلانات المتعلقة بقضايا معيّنة في الولايات المتحدة فقط. كما وأعلنت منصة Reddit على أنها ستنشئ، اعتباراً من نيسان/أبريل 2020 [208] مكتبةً تحتوي على الإعلانات السياسية في الولايات المتحدة فحسب. وأكدت أنها تحظر أي إعلان مخادع، غير صحيح، أو مُضلّل (لا الإعلانات السياسية فقط).

ولا توفّر جميع منصات الشفافية حول تنظيم المُحتوى على خدماتها. فلا تقدم شركة WeChat مثلاً أيّ إخطار بنتيجة تفعيل مصفاة المحتوى. ويبقى المُحتوى المحظور ظاهراً للشخص المُرسِل، ولكنّه لا يظهر في المُحادثة على جهاز الشخص المُستقبِل (Ruan, Knockel, Q. Ng, & Crete-Nishihata, 2016; Knockel & Xiong 2019). وهناك أيضاً غياب للشفافية على VK. ففي عام 2018، ذكرت صحيفة Tjournal أنه على الرغم من أنّ VK لا يسمح بالإعلانات ذات الطابع السياسي، تم الترويج لموقع التدوين الخاص بعمدة إحدى كبار المدن من خلال الأدوات الإعلانية التي تعتمدها هذه الشبكة الاجتماعية؛ ولكن، من جهةٍ أخرى، تم منع أحد زعماء المعارضة البارزين من نشر مُحتوى مماثل (Likhachev, 2018).

6. إشراك المُستخدمين

يتطلب إشراك المستخدمين تزويدهم بالقدرة على التحكم بالمُحتوى والحسابات والإعلانات التي يرونها. وتوفّر شركات التواصل عبر الإنترنت إمكانيات مختلفةً للمُشاركة، بما في ذلك التنبيه من مُحتوى معيّن لتتم مراجعته والقيام بتحديد الأولويات وتعليق المُحتوى لفترة معيّنة/كتمه أو حظر المُحتوى والحسابات، وتغيير فئات الإعلان التي وُضع فيها المستخدمون. والأداة الأخيرة متوفرة فقط في عدد من المنصات. ويسمح فيس بوك للمُستخدمين بتحديث «تفضيلات الإعلانات» عن طريق تغيير مجالات

....

203 https://transparency.facebook.com; https://www.facebook.com/ads/library/

204 https://transparencyreport.google.com; https://transparencyreport.google.com/political-ads/home

205 https://transparency.twitter.com/en.html; https://ads.twitter.com/transparency

206 https://www.snap.com/en-US/privacy/transparency; https://www.snap.com/en-US/political-ads/

207 https://linecorp.com/en/security/transparency/top

208 https://www.reddit.com/r/announcements/comments/g0s6tn/changes_to_reddits_political_ads_ policy/

اهتمامهم، في ما يخص المعلنين الذين يستخدمون هذه المعلومات، وخصائص الاستهداف.[209] وعلى LINE، يمكن للمستخدمين تحديد تفضيلاتهم في ما يخص المُحتوى المدعوم على إعلانات اللافتة على LINE Talk، ولكن ليس في ما يخص المُحتوى المدعوم على شريط الأحداث الجارية أو على الخدمات الأخرى (LINE, 2019a; LINE, 2019b). وكأمثلة على إشراك المستخدمين، يقدّم يوتيوب «يوتيوب للأطفال» YouTube Kids ووسائل أخرى تسمح للأهل بمتابعة وتقييد ما يُشاهده أطفالهم من مُحتوى،[210] كذلك، يسمح تويتر للمستخدمين بكتم «تغريدات» تحتوي على كلمة أو عبارة واسم مستخدم أو رمز تعبيري أو وسم «هاشتاغ» معيّن لكي لا يبرز على شريط الأحداث الجارية الخاص بهم.[211] ويقوم تويتر بتجربة دعم متخصص لما يسميه «المدافعين العاملين في الخطوط الأولى» (على سبيل المثال الصحفيين الذين يحاولون محاربة التضليل على تويتر والذين يتم استهدافهم بسبب ذلك).

7. الطعن/التظلُّم

أخيراً، ورداً على الإجراءات المتخذة في مجال تنظيم المُحتوى وتماشياً مع «مبادئ سانتا كلارا» حول الشفافية والمُساءلة في عملية تنظيم المُحتوى،[212] من المهم من وجهة نظر حماية حرية التعبير أن تضع الشركات إجراءات تسمح بالطعن في قرارات الحظر، وفي قرارات تخفيض مرتبة المُحتوى أو إزالته، وتعطيل أو تعليق الحسابات. وهذا يستلزم إخطاراً مفصّلاً بالإجراء الذي تم تطبيقه، وتوفير إمكانية للطعن في الإجراء مباشرةً من خلال خدمة الشركة، وإخطاراً بالقرار الذي تتخذه الشركة بخصوص الطعن المقدم.

وكما هو واضح من المناقشة الواردة أعلاه، تختلف الاستجابات للمعلومات المضللة. فعلى سبيل المثال، يعمد فيس بوك إلى الحد من نشر المعلومات المضللة بدلاً من إزالتها، إلا في حال كانت تنطوي أيضاً على انتهاكات أخرى للقواعد المُجتمعية (كأن تتسبب بأذىً جسديّ). وفي الوقت نفسه، وكما نوقش في القسم 4.1، لا يُخضع فيس بوك لإجراءات تنظيم المُحتوى؛ خطابات السياسيين والأحزاب السياسية والجهات التابعة لها إن كان على شكل منشورات أو إعلانات. ومن شأن هذا المزج في المعايير أن يُصعّب معالجة مسألة الطعون المتصلة بالتضليل بطريقة مباشرة. ومع ذلك، فمن الواضح أنه، على أرض الواقع، هناك مستوى من التباين في مقاربة شركات التواصل الاجتماعي لمسألة الطعون. وعلى الرغم من أنّ تقديم الطُّعون إلى هيئات تحكيمية أو قضائية ممكن من الناحية النظرية في بعض البلدان، وخصوصاً حيث يتقاطع التضليل مع القيود القانونية المحلية، يوفّر عدد قليل من الشركات فقط آليات طعن صلبة تنطبق على المحتويات والحسابات، أو لإعلام المستخدم عندما يتم اتخاذ أي إجراء.

في عام 2018، أدخل فيس بوك تعديلات على عملية تقديم التظلُّمات؛ حيث لم يكن التظلُّم متاحاً في السابق إلا في ما يخص الإجراءات المتخذة بحق الحسابات الشخصية والصفحات والمجموعات. ونتيجةً لذلك، أصبح من الممكن التظلُّم في الإجراءات المتخذة بحق المنشورات الشخصية (العري/

....

209 https://www.facebook.com/ads/preferences

210 https://support.google.com/youtubekids/answer/6172308?hl=en

211 https://help.twitter.com/en/using-twitter/advanced-twitter-mute-options

212 https://santaclaraprinciples.org/

النشـاط الجنسـي، وخطـاب الكراهيـة أو العنـف الشـديد)[213] (Bickert, 2018). علـى WeChat[214] وLINE[215]، بإمكان المسـتخدمين طلب إزالـة حجب/إلغـاء تجميد الحسـابات، ولكن لا دليل متوفر حول إمكانيـة الطعن فـي إزالـة المُحتوى. كذلـك فـإنّ الأدلـة غيـر متوفـرة لمعرفـة مـا إذا كان سـناب تشـات[216] أو واتسـاب[217] قاما بوضـع آليـة لتقديـم الطعـون. وقـد يشـكل ذلـك مشـكلةً خصوصـاً مـن منظـور حريـة التعبيـر. علـى سـبيل المثال، تقضـي إحـدى المُمارسـات المعروفـة والمسـتخدمة علـى يـد وكلاء التضليـل بالإبـلاغ الـكاذب عـن حسـابات وملفـات الصحفييـن كوسـيلة لفـرض الرقابـة عليهـم. (مُراجعـة النقـاش فـي هـذا الفصـل وفـي القسـم 7.1 حـول مجلـس الرقابـة التابـع لفيسـ بـوك).

وتتطـوّر الجهـود الهادفـة إلـى التّصـدي للمعلومـات المضللـة مـن قبـل شـركات التواصـل عبـر الإنترنـت بوتيـرة سـريعة. ولكـنّ مقاومـة هـذه الشـركات للاسـتجابة للتضليـل بشـكل ملائـم، وعلـى نطـاق عالمـي، وعـدم قيامهـا بتحمـل المسـؤولية علـى شـكل مسـؤولية الناشـر فـي مـا يخـص الآثـار الاجتماعيـة والديمقراطيـة لهـذا التضليـل؛ يضعهـا فـي خطـر أن تصبـح مصانـع لـ«اضطـراب المعلومـات» وللإسـاءة علـى الإنترنـت (Posetti, 2018b).

2.1.6 التدخّلات الصحفية لتنظيم المُحتوى:

تطبـع ميـزة التحقّـق والتدقيـق الصحافـة المهنيـة فـي جوهرهـا.[218] وقـد شـكّل تنظيـم مُحتـوى المعلومـات الوقائعيـة ونشـرها للاسـتخدام الجماعـي مـن قبـل المؤسسـات الإخباريـة، إلـى جانـب كشـف الأكاذيـب مـن خـلال صحافـة المُسـاءلة (Mayhew, 2020)، وظيفـة مهمـة تاريخيـاً لمُكافحـة التّضليـل. غيـر أن تـآكُل الوظائـف التقليديـة لحراسـة الأخبـار، إلـى جانـب «بـروز دور الجمهـور»، وانتشـار وسـائل التواصـل الاجتماعـي فـي كل مـكان؛ قوّضـت القـوة التـي كانـت تتمتّـع بهـا الأنمـاط مـا قبـل الرقميـة لتنظيـم المُحتـوى التحريـري كحصـن ضـد التّضليـل (Posetti 2018). وفـي هـذا الصـدد، كتبـت رئيسـة تحريـر صحيفـة الغارديـان «كاثريـن فاينـر» أن «فيسـ بـوك أصبـح الناشـر الأغنـى والأقـوى فـي التاريـخ مـن خـلال اسـتبدال المحرّريـن بالخوارزميـات» (Viner, 2017).

وُصفـت منصّـات التواصـل عبـر الإنترنـت بأنهـا باتـت «الحُـراس الجُـدد» للأخبـار (Bell & Owen, 2017). ومـع ذلـك، وكمـا تمـت مناقشـته فـي هـذا التقريـر البحثـي، لا تـزال هـذه المنصـات متـرددة إلـى حـدّ كبيـر فـي قبـول المسـؤولية التقليديـة المتمثّلـة فـي مراقبـة نشـر الأخبـار -بمـا فـي ذلـك التدقيـق فيهـا وتنظيـم محتواهـا- علـى الرغـم مـن اتخاذهـا قـرارات تقضـي بممارسـة الرقابـة علـى بعـض مـن المُحتـوى بطريقـة انتُقـدت بداعـي أنهـا تُقـوّض حُريّـة الإعـلام (Doyle, 2016). وقـد أُثيـر جـدل حـول حـذف معلومـات، بمـا فـي ذلـك صـور إخباريـة مهمّـة تاريخيـاً، عـلاوةً علـى تعليـق حسـابات صحفييـن لمشـاركة صـور إخباريـة يُزعـم أنهـا انتهكـت «معاييـر المجتمـع» لأنهـا صـوّرت العُـري (Kleinman, 2016; Gillespie, 2018). وقـد أطلـق عـدد مـن الموضوعـات الجدليـة هـذه -التـي نالـت تغطيـة إعلاميـة واسـعة- عمليـات أدت فـي نهايـة المطـاف إلـى إنشـاء مجلـس الرقابـة التابـع لفيسـ بـوك فـي عـام 2019.

213 https://transparency.facebook.com/community-standards-enforcement

214 https://help.wechat.com/

215 https://terms.line.me/line_terms/?lang=ja

216 https://support.snapchat.com/en-US/i-need-help?start=5153567363039232

217 https://www.whatsapp.com/legal/?eea=0#terms-of-service

218 انظر: النقاش في القسم 7.1 حول الاستجابات المعيارية والأخلاقية للمعلومات المُضلّلة.

لقد حقّق التحول الرقمي الكثير من المنافع، بما في ذلك تعزيز فرص الاستفادة من حُريّة التّعبير والوصول إلى المعلومات المتنوعة. غير أنه غذّى أيضاً تحديات وتغييرات هيكلية غير مسبوقة ومستمرة في صناعة الأخبار، تحبّذ المعلومات المُضلّلة واسعة الانتشار، بما في ذلك عن طريق تقويض الدور الذي يؤديه تنظيم المُحتوى الصحفي، وهي تشمل[219]:

- انهيار نموذج الأعمال التقليدي لنشر الأخبار، مما أدّى إلى إغلاق وسائل الإعلام بصورة متزايدة وإلى بطالة جماعية داخل القطاع، الأمر الذي قلّل بشكلٍ كبير من القدرة على تنظيم المُحتوى؛
- استنفاد موارد غرف الأخبار (من موظفين وميزانيات)، مما يؤدي إلى إبلاغ أقلّ على أرض الواقع، ويؤثر على عمليات التحقُّق من الوقائع والتحرير، ويقود بالتالي إلى تدقيق أقل في المعلومات والمصادر؛
- التقارب بين وسائل الإعلام: يُكلّف الكثير من الصحفيين في الوقت الحالي بكتابة مُحتوى لمنصات متعددة في وقت واحد (من الهاتف المحمول إلى الطباعة)، مما يقلّص بشكل إضافي الوقت المتاح لتقديم التقارير الاستباقية والتحقُّق الدقيق؛
- يُطلب من المراسلين بصورة متزايدة تحرير المُحتوى الخاص بهم ونشره دون مراجعة مناسبة؛
- زيادة الطلب على إصدار مُحتوى يغذي الصفحات الرئيسية وقنوات التواصل الاجتماعي علاوةً على ارتفاع ضغط الموعد النهائي لإنهاء المهام، بالتماشي مع انخفاض عمليات ضبط الجودة وفقدان الوظائف، مما يزيد من ضعف المعايير؛
- توقّع الجمهور الحصول على أخبار «تحت الطلب»، وتوفيرها عبر الهاتف المحمول، والمشاركة في الوقت الحقيقي على منصات التواصل الاجتماعي، الأمر الذي يزيد الضغط على اختصاصيي الأخبار الذي يواجهون تناقص الموارد في دورة الأخبار التي لا تنتهي. وتكون المواعيد النهائية للمواد المنشورة رقمياً أولاً آنية بشكل مستمر، مما يزيد من ارتكاب الأخطاء، بما في ذلك مشاركة غير مقصودة لمعلومات مضللة أو مواد من مصادر زائفة؛
- النشر «على منصات التواصل الاجتماعي أولاً» أصبح أمر شائع. حيث يقوم الصحفيون بتنظيم التغذية الإخبارية الخاصة بهم على حسابات منصات التواصل الاجتماعي لتلبية طلب الجمهور المتزايد على الأخبار في الوقت الحقيقي. وتشمل الممارسات «التغريد المباشر»، وأشرطة الفيديو المباشرة عبر فيس بوك، وغيرها من الأعمال الصحفية التي لا تنطوي بالضرورة على رقابة تحريرية (فهي أقرب إلى البث المباشر)، مما قد يؤدي إلى انتشار عقلية «النشر أولاً، والتحقُّق لاحقاً»؛
- يكافح ناشرو الأخبار من أجل التمسّك بالجماهير مع إزالة الحواجز التي تعترض النشر، الأمر الذي يتيح لأي شخص أو كيان إنتاج المُحتوى وتنظيمه، وتجاوز الحراس التقليديين، والتنافس على الاهتمام
- ويشمل هذا الأمر الجهات الفاعلة القوية التي تسعى إلى تقويض مصداقية الإبلاغ النقدي؛

[219] تُعتبر الأمثلة التالية بمثابة تنظيم للأثر وهي مستمدة من:
"News industry transformation: digital technology, social platforms and the spread of misinformation and disinformation", (Posetti, 2018), published by UNESCO: https://en.unesco.org/sites/default/files/j._jfnd_handbook_module_3.pdf

- مضايقة الصحفيين (لا سيّما النساء) ومصادرهم وجماهيرهم على الإنترنت، الأمر الذي يشتّت انتباههم ويمنعهم من التصدي للتضليل داخل مجتمعات التواصل الاجتماعي حيث يزدهر؛
- ممارسات العناوين المثيرة للانتباه clickbait (أي استخدام عناوين مضللة لإغراء القراء للنقر على الروابط تحت ذرائع كاذبة) التي صُمّمت لاستحداث حركة مرور على الإنترنت، والتي ارتبطت بتآكل الثقة في الصحافة المهنية؛
- السعي وراء الانتشار الواسع على حساب الجودة والدقة.

نتيجةً لكل ذلك، قد لا يلجأ الجمهور إلى وسائل الإعلام الإخبارية في أوقات الأزمات والكوارث وهو يتحلّى بالثقة بأنه سيحصل على معلومات جيدة التنظيم وموثوقة ومتحقق منها؛ تم نشرها ومشاركتها من أجل المصلحة العامة. وهذا الأمر قد يعرقل عملية التصدي للمعلومات المُضلّلة بشكل كبير من خلال مؤسسات متخصصة في التنظيم التحريري للمُحتوى والجماهير ومصادر المعلومات. ومع ذلك، اضطلعت بعض المؤسسات الإعلامية بتدخلات فعّالة في هذا الصدد.

من الأمثلة على ذلك، الصحافة التي تعزّز الاستجابات المتعلّقة بتنظيم المُحتوى أو تحفّزها لمُكافحة التّضليل داخل منصات التواصل الاجتماعي. وتتمثّل إحدى دراسات الحالة هذه في المقاربة التي اعتمدها الموقع الإخباري «رابلر» Rappler. فقد قام ببناء قاعدة بيانات لتتبُّع شبكات المعلومات المُضلّلة عبر الإنترنت، ثم أبلغ عن النتائج التي توصّل إليها، فأطلع منصات التواصل عبر الإنترنت على عمله. وساهمت بعض التحقيقات الجنائية الرقمية التي اضطلع بها «رابلر» Rappler في قيام فيس بوك باعتماد إجراءات لإزالة «المنشورات المُفتعلة المُعد لها سلفاً»، كما تصف الشركة حملات التّضليل المدبّرة (Rappler Research Team, 2018; Posetti et al., 2019a & 2019b; Garside, 2020).

يتجسّد مثال آخر بالتعاون في مجال التحقُّق من الوقائع بين الوسائل الإخبارية وشركات التواصل عبر الإنترنت ومنظمات التحقُّق من الوقائع وغيرهم من خبراء التدقيق الخارجيين، وهو تعاون ساعد في الحد من المعلومات المُضلّلة على وسائل التواصل الاجتماعي (انظر الوصف المفصّل لهذه المقاربات في الأقسام 4.1 و4.2 و5.3 بشأن الرصد، والتحقُّق من الوقائع، والاستجابات الاستقصائية والانتخابية). ويمكن اعتبار هذه الاستجابات استجابات تعاونية مصمّمة لتحسين تنظيم المُحتوى الخاص بوسائل التواصل الاجتماعي على مواقع/تطبيقات الشركات. فعلى سبيل المثال، قبل الانتخابات الوطنية في الهند في نيسان/أبريل 2019، أقام تطبيق واتساب شراكةً مع Proto[220]، وهي مؤسسة اجتماعية تعاونية تركّز على التحول الرقمي للصحافة، في إطار مشروع بحث عملي اسمه Checkpoint. وبموجب هذا المشروع، دُعي المستخدمون إلى الإبلاغ عن مُحتوى كاذب مشتبه فيه إلى خط مساعدة يقوم بدوره باستحداث تقارير تحقق. وبالإضافة إلى التحقُّق من المُحتوى، صُمّم هذا المشروع لجمع بيانات أبلغ عنها المستخدمون، وهي بيانات لما كانت متوفرة لولا ذلك بسبب الطبيعة المشفرة لتطبيق الدردشة «المغلق». وكان الهدف من جمع البيانات إجراء تحليل للمعلومات المُضلّلة واسعة الانتشار على واتساب، على الرغم من أنه لم يُعرف ما إذا كانت هذه التدابير أفضت إلى قيام واتساب بحظر جهات فاعلة بسبب سلوك مُفتعل مُعد له، كما يسمّيه فيس بوك.

220 https://www.checkpoint.pro.to/

3.1.6 ما ومن الذي ترصده/تستهدفه استجابات تنظيم المُحتوى؟

تركّز استجابات تنظيم المُحتوى أولاً على المُحتوى المُشارك على مواقع منصات التواصل عبر الإنترنت وتطبيقاتها، والمواد التي نشرتها الجهات الفاعلة الصحفية، ومستخدمي/جماهير كلاهما. ومع ذلك، يستخدم تطبيق واتساب (الذي يملكه فيس بوك) السلوك كوسيلة لتجنّب ممارسات مراقبة المُحتوى التي تستند إلى المُحتوى وتتطلب تعديل سياسة التشفير التام من طرف إلى طرف.[221] ويقوم الموظّفون المسؤولون عن التعلّم الآلي ومراقبة المُحتوى داخل منصّات التواصل عبر الإنترنت بالكشف عن مُحتوى يُحتمل أن يكون مسيئاً والتصرّف بشأنه، بالتعاون في كثير من الأحيان مع وكالات إخبارية. أما على المستوى الخارجي، فتساهم السُّلطات المكلّفة إنفاذ القانون، والجهات المسؤولة عن التحقُّق من الوقائع، وأطراف ثالثة أخرى أيضاً. ويتم التحقُّق من المُحتوى الذي تم وسمه في وقت لاحق أو تُنزع صفة الأولوية عنه أو تجري إزالته. وفي حالات نادرة، تنشأ أيضاً ملاحقات قضائية نتيجةً لذلك.

من حيث الأهداف، يمكن أن يكشف تنظيم المُحتوى للمستخدمين عن المُحتوى المدعوم، فيبرزه مقارنةً بما يتمّ إظهاره من خلال التشغيل العضوي لخوارزميات الشركات. وتستهدف بعض التدابير فئة محددة من المُحتوى والمُحتوى مدفوع الثمن. ومن بين التدابير التي جرى تحليلها في هذا الفصل، يستهدف عدد منها بشكل محدّد المُحتوى السياسي والجهات الفاعلة السياسية، سواءً في فترات انتخابية معينة، أو كسياسة عامة. فعلى سبيل المثال، أظهر فيس بوك/انستغرام، وغوغل/يوتيوب، وتويتر وسناب–تشات الشفافية في كيفية تنظيم الإعلان من خلال نشر «مكتبات الإعلانات السياسية»، ولكن وفق نطاق جغرافي مختلف (ومحدود في كثير من الأحيان)[222]. وفي ما يتعلق بالتحقُّق من الإعلانات السياسية، اتّخذت المنصات أيضاً خيارات مختلفة. فقد حظر تويتر الإعلانات السياسية في تشرين الثاني/نوفمبر 2019 [223]، في حين ادّعى سناب–تشات أنه يوافق على كل إعلان سياسي منشور على المنصة[224]. من جهته، قرّر فيس بوك عدم التحقُّق من فئات معينة من الإعلانات السياسية (انظر القسم 4.1)[225]، الأمر الذي قلّل من التدقيق في المعلومات السياسية المُضلّلة. أما غوغل، فقد عمد إلى تحديث سياسته لتقييد الاستهداف الدقيق السياسي (Spencer, 2019).

أما بالنسبة إلى المحتويات الأخرى المعرّضة للتضليل، مثل: الصحة والسلامة العامة، فقد لجأت المنصات الإلكترونية أيضاً إلى تكييف سياساتها. فبغرض الحد من المعلومات المُضلّلة بشأن اللقاحات على خدمات تويتر، قرّر الأخير في أيار/مايو 2019 إعادة توجيه المستخدمين إلى مصادر الصحة العامة متى كانوا يبحثون

221 يؤدي التشفير التام من طرف إلى آخر دوراً هاماً في الحفاظ على حُريّة التّعبير وحقوق الخصوصية في العصر الرقمي. وفي الواقع، اعتبر المقرر الخاص للأمم المتحدة المعني بالحق في حُريّة الرأي والتّعبير أن التشفير وعدم الكشف عن الهُويّة هما من العناصر التمكينية لحقوق الإنسان:
http://www.justsecurity.org/wp-content/uploads/2015/06/Kaye-HRC-Report-Encryption-Anonymity.pdf

222 نشر فيس بوك/انستغرام مكتبة الإعلانات السياسية في الاتحاد الأوروبي في ربيع عام 2019. تماماً كذلك، يوفر غوغل/يوتيوب تقارير شفافية حول الإعلان السياسي في الاتحاد الأوروبي والهند والولايات المتحدة. من جهته، وتماشياً مع ما تقوم به شركات أخرى، قرر سناب–تشات إطلاق مكتبة للإعلانات السياسية في الولايات المتحدة.

223 https://business.twitter.com/en/help/ads-policies/prohibited-content-policies/political-content.html

224 https://www.snap.com/en-US/ad-policies/political/

225 https://www.facebook.com/business/help/208949576550051?id=288762101909005

عن معلومات عن اللقاحات[226] (Harvey, 2019). وبصورة أكثر تحديداً، تعاونت منظمة الصحة العالمية مع منصات التواصل عبر الإنترنت للتأكُّد من تزويد المستخدمين بمعلومات موثوقة عن جائحة فيروس كورونا، في حين حرص غوغل وتويتر على أن تحتل المعلومات الصادرة عن منظمة الصحة العالمية المرتبة الأولى في الاستفسارات[227]. واستخدم فيس بوك معلومات من مصادر موثوقة، مثل: منظمة الصحة العالمية ومركز السيطرة على الأمراض والجهات التي تتحقّق من الوقائع، للحد من انتشار المعلومات المغلوطة التي تم التحقُّق منها عن الفيروس، والتزم بتقييد علامات الهاشتاغ المستخدمة لنشر المعلومات المُضلّلة حول الجائحة على انستغرام[228].

تستهدف استجابات تنظيم المُحتوى، ثانياً، الحسابات التي تُسيء استخدام شروط الخدمة الخاصة بالشركات، ومتى كان الأمر مناسباً، عندما تتعارض مع بعض الأحكام القانونية. ويمكن أن تكون هذه الجهات الفاعلة المسيئة مستخدمين فرديين، أو اختصاصيي اتصال ومُعلنين محترفين يرتكبون أعمالاً خبيثة معزولة أو مُعد لها. وفي مثل هذه الحالات، غالباً ما يتم تعطيل الحسابات أو تعليقها.

ثالثاً، يتمثّل خَيار آخر اختارته المنصات على الإنترنت في إشراك المُستخدمين في تنظيم بعض المُحتوى الذي يرونه. ويمكن القيام بذلك عن طريق إعطاء المستخدمين إمكانية وسم المُحتوى، وحظر/تعليق المُحتوى والحسابات لفترة معينة، وتغيير إعدادات التوصيات الخوارزمية، أو تغيير الفئة التي أُدرج فيها الإعلان. ويمكن أيضاً أن تُتاح للمستخدمين إمكانية الطعن في قرار مراقبة المُحتوى إذا اعتبروا أنه تمّ حظر محتواهم أو حسابهم أو تعطيله أو تعليقه بدون وجه حق.

وفي ما يتعلّق بالجهات الفاعلة الصحفية، فهي تقوم بجمع المُحتوى وتنظيمه، الأمر الذي يمكن أن يساعد منصات التواصل عبر الإنترنت في تنظيم المُحتوى، كما تبيّن من الأمثلة أعلاه، كما تقوم بخدمة جماهيرها الخاصة، التي تشمل تلك التي تُعدّ من المستجيبين التعاونيين للتضليل على مواقع منصات التواصل الاجتماعي وتطبيقاته.

وبالنسبة إلى هذه الفئة الأخيرة، يُعتبر الوصول إلى معلومات دقيقة ذات مُحتوى جيدّ التنظيم حصناً منيعاً ضد انتشار المعلومات المُضلّلة. وتشمل أهداف التنظيم الصحفي للمُحتوى أيضاً المروّجين للمعلومات المُضلّلة الذين يستغلون الأقسام المخصّصة للتعليقات في المنشورات الإخبارية وحساباتهم على وسائل التواصل الاجتماعي، إلى جانب الحسابات التي تعود لصحفيين أفراد. وتقتصر تدخلات تنظيم المُحتوى في هذه الحالات على تنظيم المُحتوى قبل مراقبته في حال التعليقات على المواقع الإخبارية، والتنظيم بعد المراقبة في حال مواقع التواصل الاجتماعي مثل فيس بوك وانستغرام. أما بالنسبة إلى تويتر وتطبيقات الدردشة، فلا تتوفّر القدرة على مُراقبة التعليقات بل تتوفّر سُلطة رصد المتابعين والحدّ من التضخيم الذي ينفّذه مستخدمون مشكوك فيهم يقومون بوسم مُحتوى أو إعادة تغريدة وإعادة توجيهه/ا. وقد سُجّلت تطوُّرات جديرة بالملاحظة في مجال تنظيم تعليقات الوكالات الإخبارية على الإنترنت، بما في ذلك اتجاه نحو إنهاء التعليقات بشكل صريح من أجل التقليل من خطاب الكراهية المرتبط بالمعلومات المُضلّلة (WAN-IFRA, 2016).

....

[226] اعتُمدت هذه السياسة في الولايات المتحدة (بالإنكليزية والإسبانية)، وكندا (بالإنكليزية والفرنسية)، والمملكة المتحدة، والبرازيل، وجمهورية كوريا، واليابان، وإندونيسيا، وسنغافورة، وفي بلدان أميركا اللاتينية الناطقة باللُّغة الإسبانية.

[227] https://twitter.com/Google_Comms/status/1222991098257108992

[228] https://about.fb.com/news/2020/05/coronavirus/

4.1.6 مَن الذي تحاول استجابات تنظيم المُحتوى مساعدته؟

تتمتّع استجابات تنظيم المُحتوى التي تطلقها منصات التواصل عبر الإنترنت بتأثير عالمي محتمل، نظراً لشيوع هذه المنصات في مختلف الأوساط الدولية. غير أن بعض التدابير باتت تُصمّم وتنفّذ بصورة متزايدة على النطاق المحلي، قبل أن يتم تنفيذها على الصعيد العالمي في بعض الأحيان بسبب تزايد الضغط الذي تمارسه الجهات الناظمة والرأي العام للرد على سياقات محلية محددة (انتخابات وأحداث كبرى وانتهاكات حقوق الإنسان وما إلى ذلك). وتُسجّل أيضاً استجابات للسميّة عبر الإنترنت تركّز على الولايات المتحدة– يُرافقها إهمال للبلدان النامية.

تنطبق هذه التدابير المتعلّقة بآليات الطعن في وسم المُحتوى ومراقبته، على سبيل المثال، على جميع مستخدمي المنصات عادةً (وفق التحديد الوارد في شروط الخدمة والمبادئ التوجيهية المجتمعية والسياسات التحريرية الخاصة بالمنصات). وتُعد بعض التدابير أكثر أهميةً بالنسبة الى السُّلطات العامة، مثل وسم سلوك غير مشروع مشتبه به، أو تعليق حسابات محددة، ويتم ذلك بموجب الالتزامات القانونية أو القواعد الخاصة بالمنضات. ومع ذلك، بالمقارنة مع جهات فاعلة أخرى، غالباً ما يُمنح الزعماء السياسيون معاملة يطبعها عدم التدخّل من الناحية العملية.

ويمكن القول أخيراً إنّ الاستجابات التي وضعتها هذه الشركات تخدم هدف الحفاظ على أنشطتها ونماذج الأعمال الخاصة بها. فتفادياً لرد فعل عنيف ولوائح تنظيمية قاسية قد تزيد من مسؤوليتها عن المُحتوى الذي نشره مستخدموها؛ من مصلحة هذه الشركات التعامل داخلياً مع قضايا التّضليل (وخطاب الكراهية) على خدماتها. ويمكن القول أيضاً إنه من مصلحة البعض الاستمرار في قبول الإعلانات السياسية المُضلّلة من منظور الربح أو المصلحة الاستراتيجية المتمثّلة في توفير ساحة تفسح المجال أمام زعماء التّضليل عن الناطقين بالحقيقة إذا كان هذا الأمر يوحي باعتماد سيناريو تنظيمي يميل الى عدم التدخّل في المستقبل.

تختلف العوامل المحفّزة وراء استجابات تنظيم المُحتوى، حسب ما إذا كانت ناتجة عن عمل طوعي اضطلعت به منصات التواصل عبر الإنترنت، أو عن ضغوط تنظيمية. وتنجم الإجراءات التي تتخذها الشركات بصورة طوعية عن افتراض يفيد بأن القواعد والمبادئ التوجيهية الواضحة للمستخدمين بشأن نشر المُحتوى، التي تتماشى من الناحية المثالية مع قواعد شفافة لمراقبة المُحتوى وأدوات التمكين، ستدفع المستخدمين إلى مقاومة المُحتوى المُضلل، بما في ذلك المُحتوى الذي ترد فيه عناصر خطاب الكراهية.

وبالمثل، تنظر الشركات في إمكانية توفير درجة ما من المراجعة الآلية للمُحتوى والحسابات، تكون ضرورية ومناسبة من أجل «تطهير» خدماتها، دون تكلفة توظيف لجان من المنظّمين البشر. وقد اقتصرت آلية عمليات مراقبة المُحتوى حتى الآن في الغالب على البريد الإلكتروني غير المرغوب فيه، والحسابات المجمّعة والآلية، وانتهاك حقوق الطبع والنشر، والمُحتوى الذي حُدّد سابقاً على أنه «مُسيء» أو «غير قانوني»، على الرغم من أن غياب الشفافية في الإبلاغ يصعّب نوعاً ما تقييم هذا الأمر. وترد هذه المسألة بالتفصيل في القسم 6.2 الذي يتناول الاستجابات التقنية/الخوارزمية.

تتبع استجابات المنصات التي تتعرّض لضغط تنظيمي من فكرة أن بعضاً من التدخل ضروري لإنفاذ القانون، علماً أن الهدف النهائي هو تهيئة بيئات تثني عن التّضليل، بما في ذلك الإساءة على الإنترنت. ويمكن أن يساعد تنظيم المُحتوى، الشركات، أيضاً على تجنّب القضايا القانونية، إذ يعزّز ثقة الجزء الأكبر من مستخدميها بأنهم في أيادي منصات تتسم «بالمواطنة الرشيدة» وتحترم التزاماً ائتمانياً يقضي برعاية مصالح عملائها والإبقاء على هذه الثقة.

أما الجهات الفاعلة الصحفية، من ناحية أُخرى، فتحفّزها إلى حدّ كبير في استجابات تنظيم المُحتوى التي تطال إدارة التعليقات ووسائل التواصل الاجتماعي الرغبة في:

- احترام مهمّتها الأخلاقية والضرورة التجارية للحصول على معلومات تم التحقُّق منها؛
- ضمان حصول جماهيرها على معلومات دقيقة وموثوقة، مع حمايتها من التعرض للتضليل المدمر؛
- حماية صحفييها ومواقعها على الإنترنت من الهجوم؛
- حماية جماهيرها من الهجوم عن طريق المعلومات المُضلّلة؛
- ضمان نزاهة صحافتها.

بالإضافة إلى ذلك، هناك حلول قائمة على التكنولوجيا لتنظيم التعليقات، كتلك التي ابتكرها مشروع كورال Coral Project (وهو في الأصل تعاون بين واشنطن بوست ونيويورك تايمز وموزيلا-فايرفوكس، ويملكه في الوقت الحالي VoxMedia[229]). (انظر: القسم التالي -6.2- لمزيد من المعلومات حول تطبيق التكنولوجيا في تنظيم المُحتوى للتصدي للمعلومات المُضلّلة).

5.1.6 ما هي المخرجات التي تنتجها استجابات تنظيم المُحتوى؟

تختلف المخرجات الناتجة عن استجابات تنظيم المُحتوى رداً على التّضليل باختلاف المقاربة المتّبعة والجهة/الجهات المعنية. ففي بعض الأحيان، تقوم شركات التواصل عبر الإنترنت أو الوكالات الإخبارية (والجهات الفاعلة الأخرى التي تنشر معلومات تخصّ المصلحة العامة) بالكشف علناً عن عدد الحسابات التي تمت إزالتها أو تعليقها، والتعليقات التي تم حذفها، والمُحتوى الذي تم تخفيض ترتيبه/الترويج له أو تصفيته أو حجبه وما إلى ذلك من خلال صفحات الشفافية أو التدوينات أو تعليقات انتقائية من ممثلين مرخّص لهم.

تختلف تقارير الشفافية الصادرة عن منصات التواصل عبر الإنترنت اختلافاً كبيراً، مما يحد من إمكانية مقارنتها. وبالمثل، فإنّ غياب التفاصيل في الإبلاغ (مثل: التعليل التفصيلي للإجراءات المتخذة) أو حتى غياب الإبلاغ عن ممارسات مراقبة المُحتوى (مثل «إلغاء» الأولوية الممنوحة لمُحتوى) يجعل من الصعب بمكان تقييم نطاق التدابير المتّخذة وفعاليتها. وعندما تنجم هذه الإجراءات عن التزامات التنظيم الذاتي التي تشرف عليها السُّلطات العامة، يجوز لها أن تنشر تقارير شفافية، كما هي الحال في إطار مدونة الممارسات المتعلّقة بالتّضليل الصادرة عن الاتحاد الأوروبي، والقانون الألماني حول شبكات التواصل

229 https://coralproject.net/about/

(انظر: القسم 5.1 حول الاستجابات التشريعية وما قبل التشريعية والسياستية). فعلى سبيل المثال، تم تغريم فيس بوك بمبلغ مليوني يورو من قبل «مكتب العدل الاتحادي الألماني» في عام 2019 لغياب الشفافية في تقاريره عن الشكاوى المقدمة والإجراءات المتخذة عند معالجة خطاب الكراهية وغيره من الجرائم الجنائية (Prager, 2019; Zeit, 2019). وأخيراً، قد تقوم أيضاً بصياغة التقارير المجالس المسؤولة عن مراقبة المُحتوى، مثلما التزم فيس بوك (2019) بالقيام به من خلال مجلس الرقابة الخاص به.

6.1.6 مَن هي الجهات الفاعلة الرئيسية في استجابات تنظيم المُحتوى، ومن يموّلها؟

تضطلع منصات التواصل عبر الإنترنت إلى حد كبير بمهمّة تنفيذ استجابات تنظيم المُحتوى على وسائل التواصل الاجتماعي بواسطة مواردها الخاصة. ومن الصعب الحصول على أرقام موثوقة عن نفقات المنصات على تنظيم المُحتوى. وعلى الرغم من أن الصورة غير مكتملة، إلا أنه بالإمكان تقديم بعض التفاصيل عن فيس بوك. فعلى سبيل المثال، ذكر موقع ذي–فيرج The Verge أن فيس بوك يقدم عقوداً بقيمة 200 مليون دولار لمتعاقدين خارجيين لمراقبة المُحتوى (Newton, 2019a). وفي الولايات المتحدة، يتقاضى متعاقدون حوالى عُشر ما يتقاضاه موظف فيس بوك العادي مقابل عمل يتسبّب للفرد بضغط نفسي كبير، يُفضي به في بعض الأحيان إلى اضطراب كرب ما بعد الصدمة. وتُعدّ الاستعانة بمصادر خارجية في جنوب شرق آسيا، وخاصة الفلبين، لمراقبة المُحتوى مسألة شائعة أيضاً بين المنصات (Dwoskin, Whalen & Cabato, 2019; Newton, 2019b).

علاوةً على ذلك، ظهرت أدلة تفيد بأن الجمع بين الإجهاد والتعرّض المتكرر (Schumaker, 2019) لنظريات المؤامرة وغيرها من المعلومات المُضلّلة قد يدفع بمراقبي المُحتوى إلى البدء بتصديق المُحتوى المغلوط الذي يُفترض بهم في الواقع أن يُراقبوه (Newton, 2019c). والأمر هذا يضع عبئاً على فيس بوك وغيرها من منصات التواصل عبر الإنترنت التي تعتمد بشكل كبير على مُراقبي المُحتوى لتنفيذ خُطوات هدفها حماية المتعاقدين معها من الآثار الضارة للتضليل، وتحسين أجورهم وظروف عملهم.

إلى جانب ذلك، خصّص فيس بوك مبلغ 130 مليون دولار لتشغيل مجلس الرقابة الخاص به على مدى السنوات السّت المقبلة (Harris, 2019). وأخيراً، في إطار مشروع الصحافة على فيس بوك Facebook Journalism Project، أعلن فيس بوك أنه سيطلق صندوقاً بقيمة مليون دولار لدعم التحقُّق من الوقائع[230] وصندوقاً بقيمة مليون دولار لدعم التقارير الإخبارية[231] عن جائحة فيروس كورونا. وبالمثل، سيؤمن تويتر مليون دولار لحماية الصحفيين ودعمهم خلال جائحة فيروس كورونا (Gadde, 2020). يُشار في هذا الصدد إلى أن الإيرادات السنوية لفيس بوك بلغت 70.7 مليار دولار في عام 2019 [232]. وبما أن نموذج الأعمال الخاص بالشركات يعتمد بشكل رئيسي على الإعلانات الهادفة، بإمكان المرء أن يعتبر أنه نظراً إلى أن هذا المخطط الإعلاني يستند إلى بيانات تم جمعها من المستخدمين، فإنّ المستخدمين هم الذين يُموّلون بشكلٍ غير مباشر تكاليف الاستجابات للتضليل.

[230] https://www.facebook.com/journalismproject/coronavirus-grants-fact-checking

[231] https://www.facebook.com/journalismproject/programs/community-network/coronavirus-grants-news-reporting

[232] https://www.statista.com/statistics/268604/annual-revenue-of-facebook/

في حال استجابات تنظيم المُحتوى رداً على التّضليل التي تُطلقها الجهات الفاعلة الصحفية، يتم تمويلها إما من قبل الوكالات الإخبارية نفسها، أو من قبل صحفيين أفراد يعملون بشكل مستقل لإدارة حساباتهم على وسائل التواصل الاجتماعي، أو عن طريق منح من مؤسسات أو منصات تواصل عبر الإنترنت تم تصميمها لتحسين تنظيم المُحتوى من جانب الجمهور والإدارة المجتمعية.

7.1.6 دراسة حالة بشأن الاستجابة: معلومات مُضلّلة عن جائحة فيروس كورونا

أ. استجابات منصات التواصل عبر الإنترنت

سُجّلت ردود أفعال غير مسبوقة على «مشكلة التّضليل» من قبل منصات التواصل عبر الإنترنت من أجل الحد من انتشار معلومات مغلوطة تتعلّق بالصحة وإعادة توجيه المستخدمين نحو مصادر موثوقة (Posetti & Bontcheva, 2020a; Posetti & Bontcheva,2020b)، وشملت التدابير تنفيذاً أكثر صرامة لسياساتها واعتماد إجراءات طوارئ، إلى جانب توسيع نطاق تطبيق السياسات لتشمل الجهات الفاعلة السياسية في بعض الحالات. وقد دفع هذا الوضع الفريد الشركات إلى العمل بشكل وثيق مع بعضها البعض. وقد قامت حتى بنشر بيان مُشترك اعتمده فيس بوك، وغوغل، ولينكدن، وميكروسوفت، ورديت Reddit، وتويتر، ويوتيوب، في خطوة لمُكافحة الغش والتّضليل على خدماتها[233].

ولأغراض هذا التقرير البحثي، درسنا التدابير التي اتخذتها هذه الشركات. وقد أخذ كل منها زمام المبادرة لإعادة توجيه المستخدمين نحو معلومات موثوقة والحد من انتشار المعلومات المُضلّلة. وجاءت بعض هذه التدابير استباقية، في حين اتُّخذت تدابير أخرى بعد مناقشة أُجريت مع السُّلطات العامة. في الجدول والنص أدناه، يرد تحليل إضافي شمل عدداً من أكبر منصات التواصل عبر الإنترنت. وتشير علامتا XX إلى المجالات التي اتخذت فيها المنصات الإلكترونية تدابير إضافية للحد من انتشار المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا.

---

[233] https://twitter.com/fbnewsroom/status/1239703497479614466?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1239703497479614466&ref_url=https%3A%2F%2Ftechcrunch.com%2F2020%2F03%2F16%2Ffacebook-reddit-google-linkedinmicrosoft-twitter-and-youtube-issue-joint-statement-on-misinformation%2F

مراقبة المُحتوى/ الحساب		فيس بوك/انستغرام	واتساب	غوغل يوتيوب	تويتر
وسم المُحتوى ومراجعته	مسيّر بالآلة	XX		XX	XX
	يسيره الإنسان	X		X	X
	مراجعة أجرتها أطراف ثالثة	XX	X	X	X
	مشورة خارجية	X		X	
مصفاة المُحتوى وإزالته وحجبه وتقييده	إعادة تحميل المصفاة	X		X	X
	تقييد تحويل المُحتوى		X		
	قيود استناداً إلى: قواعد المنصة	X	X	X	XX
	إنفاذ القانون	X	X	X	X
	إزالة المعلومات المُضلّلة بشكل استباقي	XX		XX	
الترويج للمُحتوى وخفض ترتيبه	الترويج لمصادر موثوقة	XX		XX	X
	خفض ترتيب العناوين المثيرة للانتباه clickbait	X		X	X
تعطيل الحسابات وتعلقيها	مقاربة متدرّجة: تحذير			X	
	• ميزات محدودة	X	X	X	X
	• تعليق	X	X	X	X
الشفافية في المُحتوى المدعوم	تحديد المُحتوى المدعوم	X		X	X
	مركز شفافية الإعلانات	X		X	X
	إزالة الإعلانات التي تشدّد على الأزمة	XX		XX	
تمكين المستخدمين	يستطيع المستخدم وسم المضمون لمراجعته	X	X	X	X
	يستطيع المستخدم حجب/تعليق المُحتوى/ الحسابات لفترة معينة	X	X	X	X
	يستطيع المستخدم منح الأولوية للمُحتوى/ الحسابات	X		X	X
	يستطيع المستخدم تغيير فئة الإعلانات التي يرد فيها	X		X	X
الطعن/التظلُّم	إنذار بالإجراء	X		X	X
	إمكانية الطعن	X		X	X
	إخطار بقرار الطعن	X		X	X

الجدول 6. استجابات تهدف الى تنظيم المُحتوى من قبل منصات التواصل عبر الإنترنت في ما يخصّ أزمة المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا

1. وسم المُحتوى ومُراجعته

بالإضافة إلى الشراكات مع الجهات المعنية بالتحقُّق من الوقائع، نفّذت عدة منصات تدابير إضافية لإزالة المُحتوى الذي وسمته سلطات الصحة العامة خلال الجائحة.

للحد من انتشار فيروس كورونا، شجّعت منصات التواصل عبر الإنترنت والسُّلطات الحكومية حجر العمّال في المنازل. ومع ارتفاع عدد الموظفين الذين يعملون عن بُعد، اختارت الشركات الاعتماد بشكل متزايد على الخوارزميات لمراقبة المُحتوى. وكانت هذه حال فيس بوك/انستغرام (Jin, 2020)، ولكن أيضاً تويتر (Gadde & Derella, 2020) وغوغل/يوتيوب (Pichai, 2020). وتماماً كما توقّعت المنصات، أدت الزيادة في المراقبة الآلية للكثير من الأخطاء و«الإيجابيات الكاذبة»[234].

2. تفعيل مصفاة المُحتوى وإزالته وحجبه وتقييده

للحد من نشر السرديات المُضلّلة بشأن فيروس كورونا، اعتمدت الكثير من هذه المنصات أيضاً مقاربة استباقية بصورة إضافية لإزالة المُحتوى. وقد أشار غوغل إلى إزالة معلومات مضللة بشكل استباقي من خدماته، بما في ذلك يوتيوب وخرائط غوغل Google Maps. فعلى سبيل المثال، أزال يوتيوب مقاطع فيديو تروّج لعلاجات غير مثبتة طبياً (Pichai, 2020). والتزم فيس بوك بإزالة «ادعاءات متعلقة بعلاجات أو أساليب وقاية زائفة –كأن يؤدي تناوُل مواد التبييض إلى الشفاء من فيروس كورونا – أو الادعاءات التي تستحدث الارتباك حول الموارد الصحية المتوفرة» (Jin, 2020). والتزمت الشركة أيضاً بإزالة علامات الوسم «هاشتاغ» المستخدمة لنشر معلومات مضللة على انستغرام. أما تويتر، فقد وسّع تعريف الأذى على المنصة ليشمل إنكار توصيات سلطات الصحة العامة، ووصف علاج معروف بأنه غير فعّال، وإنكار الحقائق العلمية حول انتقال الفيروس، وادعاءات بأن جائحة فيروس كورونا جزء من مؤامرة للتلاعب بالناس، والتحريض على إجراءات يمكن أن تسبّب الذعر على نطاق واسع، أو ادعاءات بأن مجموعة معينة ستكون أكثر عُرضةً لفيروس كورونا أو لن تكون أبداً عُرضةً له.

3. الترويج للمُحتوى وخفض ترتيبه وإشراك المستخدمين

قضت الاستراتيجيات الأساسية لمنصات التواصل عبر الإنترنت الآيلة إلى مواجهة المعلومات المُضلّلة المتعلقة بجائحة فيروس كورونا بإعادة توجيه المستخدمين نحو معلومات من مصادر موثوقة، لا سيّما من خلال ميزات البحث الخاصة بمنصات الكيانات المختلفة، والترويج لمُحتوى موثوق على الصفحات الرئيسية، ومن خلال مجموعات مخصصة لهذا الغرض. فعلى فيس بوك وانستغرام (Jin,2020)، أظهرت عمليات البحث عن علامات الوسم الخاصة بالفيروس نوافذ منبثقة تعليمية، وجرت إعادة توجيه المستخدم إلى معلومات صادرة عن منظمة الصحة العالمية والسُّلطات الصحية المحلية. كما مُنحت منظمة الصحة العالمية وغيرها من المنظمات إعلانات مجانية من قبل الكثير من شركات التواصل عبر الإنترنت لتنظيم حملات إعلامية وتثقيفية. وسلّط غوغل الضوء أيضاً على المُحتوى الذي تنشره مصادر موثوقة عندما كان الأشخاص يبحثون عن معلومات حول فيروس كورونا، بالإضافة إلى لوحات معلومات لإضافة سياق إضافي. وظهرت على الصفحة الرئيسية على يوتيوب مقاطع فيديو أعدّتها وكالات الصحة العامة (Pichai, 2020). وبشكل مماثل، عندما بحث المستخدمون عن جائحة فيروس كورونا على منصّة تيك–توك، عُرضت عليهم لافتة معلومات

[234] https://twitter.com/guyro/status/1240063821974138881

لمنظمة الصحة العالمية (Kelly, 2020a). وفي الوقت نفسه، قام تويتر بتنظيم مُحتوى صفحة أحداث خُصّصت لجائحة فيروس كورونا تعرض أحدث المعلومات من مصادر موثوقة لتظهر أعلى الخط الزمني (Gadde & Derella, 2020). واستخدم سناب–تشات، من جهته، وظيفة «ديسكوفري Discovery» لتسليط الضوء على معلومات واردة من الشركاء (Snapchat, 2020).

4. تعطيل الحسابات وتعليقها

لم تنفّذ منصات التواصل تدابير إضافية في ما يتعلق بتعطيل الحسابات وتعليقها رداً على المعلومات المُضلّلة بشأن جائحة فيروس كورونا. ومع ذلك، عمل تويتر على التحقُّق من حسابات تحتوي على عناوين بريد إلكتروني من مؤسسات صحية للإشارة إلى معلومات موثوقة حول الموضوع[235].

5. الشفافية في مراقبة المُحتوى والمُحتوى المدعوم

مُنحت منظمة الصحة العالمية وغيرها من المنظمات الموثوقة إعلانات مجانية من قبل فيس بوك، وتلقّت مساعدة للإعلان من غوغل. وفي ما يتعلق بالمُحتوى المدعوم، اختارت معظم المنصات حظر الإعلانات التي تحاول الاستفادة من الوباء. ومع ذلك، ظهرت عمليات احتيال كثيرة على وسائل التواصل الاجتماعي، مما دفع بسلطات إنفاذ القانون وحماية المستهلكين إلى تحذير المستهلكين ودعوة هذه الوسائل إلى الردّ بسرعة[236].

6. الطعن/التظلُّم

لم تُلاحظ أي تغييرات محددة في آليات التظلم في مجال فيروس كورونا، على الرغم من أن جائحة فيروس كورونا أدت إلى استنفاد القوى العاملة وزيادة الاعتماد على المراقبة الآلية للمُحتوى المضلل حول جائحة فيروس كورونا. وقد حذّر فيس بوك من احتمال حدوث المزيد من الأخطاء وأشار إلى أنه لم يعد بإمكانه ضمان وصول المستخدمين الذين طعنوا بالإزالة التلقائية للمُحتوى إلى عملية مراجعة يجريها الإنسان. وصدرت بيانات مماثلة عن غوغل وتويتر ويوتيوب. وفي الحالات التي أخطأ فيها الكشف والتصفية الآلية (على سبيل المثال، تمت إزالة منشور لمستخدم مرتبط بأخبار أو مواقع إلكترونية مشروعة خاصة بجائحة فيروس كورونا)، تسبّب تمييع حق الطعن وغياب آلية تصحيح قوية بضرر محتمل طال حقوق المستخدمين في حُريّة التّعبير (Posetti& Bontcheva, 2020a). وقد أضعف هذا الأمر أحد الالتزامات المؤسسية الرئيسية التي ركّز عليها مقرر الأمم المتحدة الخاص المعني بالحق في حُريّة الرأي والتّعبير (UN Special Rapporteur on Freedom of Opinion and Expression, 2018b, section IV, pars 44-63).

ب. استجابات تنظيم المُحتوى رداً على «مشكلة التّضليل» التي أطلقتها الجهات الفاعلة الصحفية

اعتُبرت استجابات تنظيم المُحتوى أيضاً ميزة رئيسية في استراتيجيات الوكالات الإخبارية لمُكافحة «مشكلة التّضليل» (Posetti & Bontcheva, 2020a). فإلى جانب تشديد الرقابة على التعليقات على الإنترنت وزيادة الوعي حول المخاطر المتزايدة الناجمة عن انخراط الجمهور على قنوات التواصل الاجتماعي ذات العلامات التجارية، مثل فيس بوك، حيث لا يمكن مراقبة التعليقات بصورة مُسبقة، أطلق ناشرو الأخبار منتجات تحريرية منظّمة المُحتوى ومصمّمة خصيصاً لتثقيف جماهيرهم وإعلامها.

[235] https://twitter.com/twittersupport/status/1241155701822476288?s=12

[236] https://www.consumer.ftc.gov/features/coronavirus-scams-what-ftc-doing

من الأمثلة على هذه التدخلات الصحفية الهادفة إلى تنظيم المُحتوى ما يلي:

- النشرات الإخبارية المواضيعية التي تنظّم مُحتوى أفضل التقارير والأبحاث وعمليات كشف الزيف على أساس جدول زمني محدّد[237].
- المدونات الصوتية (بودكاست) التي تدحض الخرافات من خلال تنظيم مُحتوى التحقُّق من الوقائع والمقابلات ومراجعة البيانات ومعلومات الصحة العامة الموثوق بها حول جائحة فيروس كورونا[238].
- المدونات الحية[239] والقوائم[240] وقواعد البيانات المحدّثة بانتظام بشأن معلومات مضّللة كُشف زيفها من جميع أنحاء العالم[241].
- عمليات متخصصة لتنظيم المُحتوى تجمع بشكل مركزي الموارد والمبادئ التوجيهية والتقارير التفسيرية حول ممارسة الصحافة بأمان وبأخلاقية وفعالية خلال الوباء[242].

بالإضافة إلى ذلك، جمعت المنظمة غير الحكومية First Draft قائمة بالأساليب التي اعتمدتها 11 منصة إنترنت رئيسية للاستجابة لما سُمّي بالمعلومات المغلوطة والمُضلّلة حول جائحة فيروس كورونا[243]. وشملت بعض الإجراءات الرئيسية التي تم تحديدها إلغاء تسجيل المروّجين الواضحين للمعلومات المُضلّلة، مع رفع مستوى المصادر الموثوقة من خلال توفير حيز إعلاني مجاني وآليات أخرى. بما أنّ المؤسسات الإعلامية هي الحارس التقليدي في مجال إنتاج المُحتوى ونقله، فهي تواجه تحديات خاصة تتعلق «بمشكلة التّضليل». في الواقع، يُعدّ التنوع الإعلامي مساهمة قيّمة في المجتمع، غير أن بعض ناشري الأخبار باتوا تحت سيطرة قوى تقوم بتسييس الأزمة بشكل لا مبرر له؛ يُرقى إلى مستوى التّضليل. نتيجةً لذلك، يُعتبر بعض الصحفيين معرّضين للخدع والإثارة والممارسة الإشكالية من الناحية الأخلاقية المتمثلة في تفسير الالتزام بالموضوعية بشكل خاطئ من خلال مقاربة «التوازن الخاطئ»، حيث يقيّمون المصادر غير الصادقة والصادقة بالتساوي، وفي كثير من الأحيان، بشكل غير نقدي (Posetti & Bontcheva, 2020b). وقد أدت هذه الظواهر إلى قيام بعض وسائل الإعلام الإخبارية بإضفاء الشرعية على معلومات مضللة حول جائحة فيروس كورونا (Moore, 2020; Henderson, 2020). يتعارض هذا الفشل في النظام مع دور الصحافة كعلاج للتضليل، كما أنه يقلّل من قدرة وسائل الإعلام الإخبارية على لفت الانتباه إلى فشل النظام على نطاق أوسع، كالافتقار إلى المعلومات الرسمية وجهوزية الموارد العامة أو سوء توجيهها.

---

237 انظر، على سبيل المثال، the Infodemic Newsletter الصادرة عن CodaStory
https://mailchi.mp/codastory/the-infodemic-may-3726181?e=57d6fdb385

238 انظر على سبيل المثال ABC Australia’s ‘Coronacast’ podcast
https://www.abc.net.au/radio/programs/coronacast/

239 انظر، على سبيل المثال، التدوين المباشر الشامل عن الوباء الخاص بالغارديان The Guardian
https://www.theguardian.com/world/live/2020/mar/31/coronavirus-live-news-usa-confirmed-cases-double-china-update-uk-italy-spain-europe-latest-updates

240 انظر تنظيم Buzzfeed لمُحتوى الأساطير والخدع عن فيروس كورونا COVID-19
https://www.buzzfeednews.com/article/janelytvynenko/coronavirus-fake-news-disinformation-rumors-hoaxes

241 انظر تنظيم معهد بوينتر Poynter لمُحتوى التحقُّق من الوقائع وكشف الزيف حول جائحة فيروس كورونا
https://www.poynter.org/fact-checking/2020/the-coronavirusfacts-global-database-has-doubled-in-a-week-check-out-the-latest-hoaxes-about-covid-19/

242 انظر الموارد المنظمة للمركز الدولي للصحافة للمساعدة في الإبلاغ عن جائحة فيروس كورونا،
https://ijnet.org/en/stories#story:7100

243 https://firstdraftnews.org/latest/how-social-media-platforms-are-responding-to-the-coronavirus-infodemic/

شكّلت جائحة فيروس كورونا فرصةً بالنسبة إلى الكثير من ناشري الأخبار والصحفيين لتعزيز خدمتهم العامة من خلال تعزيز استقلالية التحرير والالتزام بأعلى معايير الأخلاقيات والمهنية؛ تماشياً مع آليات قوية للتنظيم الذاتي. وتمكّنت الصحافة بهذه الطريقة من إثبات مُساءلتها عن المعايير، وتمييز نفسها عن المُحتوى الإشكالي والتفاعل السائد في فضاء الرسائل الخاصة والمباشرة الآخذ في الاتساع (بما في ذلك تطبيقات الرسائل، مثل: واتساب)، حيث تزدهر المعلومات المُضلّلة ووكلاؤها بعيداً عن أنظار الجمهور الأوسع نطاقاً، وتستمر دون رادع. وقد تمكّن ناشرو الأخبار من إظهار جدارتهم بالثقة كمصدر للوقائع والرأي القائم على الوقائع، وقاموا بتعزيزها من خلال الكشف عن الجهات الفاعلة المنظّمة في إطار «مشكلة التّضليل». وبالمثل، سلّطوا الضوء على دورهم الهام في ضمان الاستجابات الشفّافة الخاضعة للمُساءلة العامة من قبل جميع الجهات الفاعلة إزاء كل من «مشكلة التّضليل» وجائحة فيروس كورونا الأوسع نطاقاً.

8.1.6 كيف يتم تقييم هذه الاستجابات؟

تتألّف الاستجابات الهادفة إلى تنظيم المُحتوى التي وضعتها شركات التواصل عبر الإنترنت في المقام الأول من تدابير التنظيم الذاتي، وهي بالتالي لا تتبع هيكلاً متسقاً للإبلاغ. وتعطي المبادئ التوجيهية وتقارير الشفافية وتدوينات المنصات أو الإعلانات العرضية فكرة بدائية عن عمليات صنع القرار في المنصات المعنية. ويشير التقييم الذي أجرته الحكومات[244]، والوسط الأكاديمي (Andreou et al., 2018)، والإعلام (Lomas, 2020)، ومجموعات المجتمع المدني (Privacy International, 2020) إلى كل من قيمة تنظيم المُحتوى الذي تضطلع به منصات التواصل على الإنترنت وإلى حدوده المحتملة (انظر: المناقشة أعلاه حول «مبادئ سانتا كلارا»، وأدناه بالنسبة إلى المقرر الخاص للأمم المتحدة في القسم المعنون «التحديات والفرص»). وفي بعض الحالات، تقيّم الجهات الناظمة أيضاً التزامات التنظيم الذاتي بهدف وضع مقترحات تنظيمية جديدة محتملة. فعلى سبيل المثال، تتضمّن مدونة ممارسات الاتحاد الأوروبي بشأن التّضليل تقييماً لالتزامات المفوضية الأوروبية والجهات الناظمة، قبل أي تنقيح أو اقتراح تنظيمي محتمل.

وفي الحالات التي صدرت فيها تشريعات تلزم المنصات الإلكترونية بالتصدي لانتشار المعلومات المُضلّلة (انظر القسم 5.1 حول الاستجابات التشريعية وما قبل التشريعية والسياستية)، يمكن إدراج معايير التقييم على نحو أكثر انتظاماً. فعلى سبيل المثال، في شباط/فبراير 2020، وافقت الحكومة الألمانية على حُزمة تنظيمية لتحديث القانون المتعلّق بشبكات التواصل لعام 2017 وإكماله (German BMJV, 2020a; German BMJV, 2020b).

وبالنسبة إلى قيام الجهات الفاعلة الصحفية بتقييم استجابات تنظيم المُحتوى رداً على المعلومات المُضلّلة، لا تتوفّر عملية تقييم منهجية بل تنطبق مجموعة متنوعة من التدابير، تتراوح ما بين نشاط الصحفيين الفرديين وعمليات مُراجعة النُّظراء (Peer Review)، مثل مجالس الصحافة والجوائز المهنية. فعلى مستوى الصحفيين والمؤسسات الإخبارية الفردية، تقيس مقاييس التواصل الاجتماعي وتحليلات

[244] انظر: على سبيل المثال، محاولات البرلمان البريطاني للتدقيق في المقاربات التي اعتمدتها منصات التواصل عبر الإنترنت لتنظيم المعلومات المُضلّلة خلال جائحة فيروس كورونا:
https://www.parliament.uk/business/committees/committees-a-z/commons-select/digital-culture-media-and-sport-committee/sub-committee-on-online-harms-and-disinformation/news/misinformationcovid-19-19-21/

غرف الأخبار بعض نتائج تنظيم المُحتوى، بما في ذلك نطاق إشراك الجمهور و«مثابرته» (مثلاً، الوقت الذي يقضيه المرء في قراءة مقالة معينة، وعدد الاشتراكات/الأعضاء الجدد، والمتابعة، والمشاركة والتعليقات). والأمر هذا لا يعطي بالضرورة انطباعاً دقيقاً عن الأثر، لأن القصص أو التدوينات ذات الجمهور المنخفض نسبياً قد تستمر في تحقيق أثر كبير على صعيد السياسة العامة على المستوى القُطري أو المستوى الحكومي الدولي.

على ذلك، تعترف الجوائز المهنية بدور التدخلات التحريرية في أزمة المعلومات المُضلّلة. فعلى سبيل المثال، نال الفائزون المشتركون أكبر جائزة دولية للصحافة الاستقصائية في عام 2019 (جائزة Global Investigative Journalism Network's Shining Light Award) استناداً إلى سلسلة من التقارير ومُحتوى آخر خضع للتنظيم وساعد على كشف شبكات من المعلومات المُضلّلة مرتبطة بالدولة في جنوب أفريقيا والفلبين (Haffajee, 2019).

9.1.6 التحديات والفرص

أوضحت حملات التّضليل السابقة أنه بدون تدخّل لتنظيم المُحتوى، سيصعُب التنقّل بين الخدمات التي تشغّلها منصات التواصل عبر الإنترنت واستخدامها بسبب سيول الرسائل غير المرغوب فيها، والمُحتوى المسيء وغير القانوني، والمستخدمين الذين لم يتم التحقُّق منهم. وبما أن الشركات نفسها تستطيع الوصول إلى بيانات عن مستخدميها، فهي مخوّلة برصد المُحتوى ومراقبته وفقاً لسياساتها وتقنياتها. وقد يساعد وضع استراتيجيات، مثل حظر ما تصفه المنصات أحياناً «بالسلوك المُفتعل المُعد له» من خدماتها، أو الترويج للمُحتوى الذي تم التحقُّق منه، في الحد من انتشار المُحتوى المغلوط والمضلل، والسلوكيات المسيئة المرتبطة به. ومع ذلك، توضع السياسات بالطريقة الفضلى من خلال عمليات تشمل أصحاب مصلحة متعددين، ويتعيّن بالتالي تنفيذ هذه السياسات على نحو متّسق وشفاف. ويمكن أيضاً المساعدة في رصد هذا التنفيذ من خلال تمكين الباحثين الممتثلين أخلاقياً من الوصول بشكل أكبر إلى بيانات الشركات.

وتتيح مقاربة تحبذ التعاون والانخراط مع أصحاب مصلحة آخرين، بما في ذلك الجهات التي تتحقّق من الوقائع والمجالس الاستشارية المستقلة، إجراء رقابة خارجية. وتتيح هذه المقاربة أيضاً تفادي التدخلات القانونية التي يمكن أن تحد من دون مبرر من حُريّة التّعبير. وتتماشى هذه المقاربة مع وجهة نظر مؤتمر القمة العالمي لمجتمع المعلومات، الذي يحثّ أصحاب المصلحة المتعددين على الانخراط في مسائل الحوكمة، بدءاً من المبادئ ووصولاً إلى القواعد التشغيلية (World Summit Working Group, 2005).

ويصعب رصد فعالية الاستجابات الرامية إلى تنظيم المُحتوى وتقييمها في غياب إفصاح أكبر من قبل منصات التواصل عبر الإنترنت. وقد أدى هذا الأمر إلى جدل متزايد تمحور حول قيام الشركات بتحديد المُحتوى والحسابات التي تنشر المعلومات المُضلّلة وتوزّعها وتخفض ترتيبها وتحذفها. وفي موازاة ذلك، يبرز قلق بشأن استثناءات خاصة من هذه القواعد تستفيد منها شخصيات سياسية قوية.[245] فعلى سبيل

[245] انظر النقاش السابق في هذا الفصل بشأن تويتر وفيس بوك والرئيس الأميركي السابق، إلى جانب تحليل للجدل هذا في القسمَين 5.3 و7.1.

المثال، لا يتّضح كم مرة أو في أي ظرف تتم تصفية المُحتوى والحسابات وحجبها مسبقاً على منصات هذه الكيانات. وتقود الآلة بعض عمليات المراجعة والرصد، استناداً إلى مسح قواعد «بيانات الهاش التقنية» hash databases وأنماط السلوك المُفتعل. ولكن لا يتّضح ما هي الضمانات المتوفّرة لمنع الإفراط في تقييد المُحتوى والحسابات[246]. تؤكّد صحة هذا القول خلافات متصلة بحذف غير مناسب تم تبريره على أساس خرق قواعد المنصة. وتستطيع الاستجابات الرامية إلى تنظيم المُحتوى أن تؤدي إلى الكثير من الإيجابيات/السلبيات الخاطئة[247]، لا سيما عندما تكون آلية لا وجود للعنصر البشري فيها. ويُلاحظ غياب للشفافية بشأن تواتر التصفية وفئاتها، فيما تكون آليات الطعن في استجابات تنظيم المُحتوى ضعيفة بشكل عام في معظم المنصات. وتثير كل هذه المسائل مجتمعةً شواغل رئيسية من منظور حُريّة التّعبير.

يمكن اتخاذ إجراءات الإنصاف على أساس القانون ومعايير المجتمع القائمة. ومع ذلك، لابدّ من التنبّه إلى وجود قيود لفائدة امتثال منصات التواصل الاجتماعي للوائح التنظيمية الوطنية، لأنها تعمل على الصعيد العالمي ولا يقع نطاق عملها بالضرورة ضمن الأُطُر القانونية للولايات القضائية التي تعمل فيها. وتفضّل الشركات العمل على نطاق واسع من حيث القانون: فهي عادةً ما تتّخذ من ولاية قضائية واحدة مقراً لها، ولكن مستخدميها يعبرون الولايات القضائية. ويُعتبر التقيّد بالقوانين الوطنية غير متكافئ. وفي بعض الحالات، تتبع سياسات مراقبة المُحتوى ومعاييرها تفسير المقرّ لمعايير حُريّة التّعبير، على نحو أوثق من نظام وطني معين. وفي بعض الحالات، يفيد هذا الأمر المستخدمين كأولئك الموجودين في ولايات قضائية تكون فيها القيود أدنى من المعايير الدولية التي تحدّد أي خطاب يتمتع بالحماية.

في الوقت نفسه، غالباً ما تميل شروط الخدمة والمبادئ التوجيهية المجتمعية وسياسات التحرير إلى التحلي بطابع تقييدي أكبر، فتحدّ بالتالي من الكلام، بما يتجاوز ما هو مطلوب قانوناً على الأقل في الولاية القضائية التي حصل فيها التسجيل القانوني (مثل: الرقابة التي يفرضها فيس بوك على العري). وبالتالي، تحدّد الشركات الخاصة ذات النطاق العالمي إلى حد كبير، وبطريقة غير منسقة حالياً، ما هو التّعبير المقبول في إطار إنفاذ معاييرها الخاصة. ويمكن أن يؤدي ذلك إلى تصرّف هذه الشركات كجهات تحدّد حُريّة التّعبير على خدماتها وتحكم عليها وتنفّذها. وفي الواقع، باستطاعة أي تحرك تقوم به هذه الشركات في جانب المراجعة ومراقبة المُحتوى والشفافية وإشراك المستخدمين، والطعن/التظلّم من القرارات أن يترك آثاراً سلبية هائلة على حُريّة التّعبير.

وتزداد الأمور تعقيداً لأنه، مع الاعتراف بالدور الذي يتعيّن على منصات التواصل عبر الإنترنت أن تؤديه في الحد من المعلومات المُضلّلة المنشورة على صفحاتها، قد تبرز مشكلات بسبب تفويض الدول غير الرسمي للسلطة التنظيمية لهذه المنصات الخاصة. والأمر يصحّ بصفة خاصة عندما يقلل هكذا تفويض من المساءلة والصلاحية القضائية للقرارات المتعلقة بحُريّة التّعبير عموماً، والتي تُعد من مسؤوليات الدول وينبغي عليها بالتالي أن تمتثل للمعايير الدولية لحقوق الإنسان. وقد يُرقى هذا الأمر الى مستوى الرقابة المخصخصة. ومتى نصّت اللوائح التنظيمية صراحةً على التفويض (انظر: القسم 5.1 الذي يتناول الاستجابات التشريعية وما قبل التشريعية والسياستية)، يمكن أن تتوفّر مساءلة عامة

---

246 انظر المزيد من التفاصيل في القسم 6.2 – الاستجابات التقنية/الخوارزمية.

247 في سياق جائحة فيروس كورونا، عزّز فيس بوك مراقبته للمُحتوى المتعلّق بهذه المسألة. غير أن استخدام المصفاة الآلية ضد الرسائل الإلكترونية غير المرغوب فيها أدى إلى إزالة مصادر ذات مصداقية. https://twitter.com/guyro/status/1240063821974138881

عن هذه اللوائح التنظيمية في الديمقراطيات التي تحترم سيادة القانون ومسألتَي الضرورة والتناسب. ومع ذلك، تجد شركات الإنترنت نفسها في الوقت ذاته مسؤولة إلى حد كبير عن التنظيم الذاتي للمُحتوى كما تراه مناسباً، بحكم الواقع ولأسباب سياسية واقتصادية وتكنولوجية مختلفة.

هناك جانب من الصحة في المخاوف المرتبطة بحُريّة التّعبير التي تعتبر أن تنظيم المُحتوى وفق اللوائح التنظيمية يمكن أن يكون أسوأ من التنظيم الذاتي للمُحتوى في أجزاء مختلفة من العالم. على الرغم من ذلك، فإن تنظيم المُحتوى ذاتياً يعرّض عموماً للمسؤولية القانونية بموجب القوانين المتعلقة بحقوق التأليف والنشر وإساءة معاملة الأطفال، على سبيل المثال. لذلك، تتعلّق المسألة بأنواع التنظيم بدلاً من التنظيم بحد ذاته. وتزداد الأمور صعوبةً عندما تجرّم اللوائح التنظيمية المعلومات المُضلّلة، لا سيما عندما تكون غامضة و/أو غير متناسبة من حيث المعايير الدولية لحقوق الإنسان. ومع ذلك، يمكن أن تكون اللوائح التنظيمية للمستهلك بشأن حماية البيانات والقدرة على الطعن في القرارات، فضلاً عن اللوائح التنظيمية الخاصة بالشفافية التي تطلب من الشركات الإبلاغ عن كيفية اتخاذ القرارات، أقل تعقيداً من منظور حُريّة التّعبير.

وكما هو مبيّن في مقدمة هذا الفصل، تقدم كل شركة من شركات التواصل عبر الإنترنت أنواعاً مختلفة من الخدمات وتعمل بطرق مختلفة، مما يبرر الحاجة إلى التفريق في القواعد المتعلقة باستخدام خدماتها. غير أنه في غياب معايير وتعاريف متفق عليها، تستخدم كل شركة «مقياس تنظيم المُحتوى» الخاص بها، دون أي اتساق في الإنفاذ أو الشفافية أو الطعن عبر المنصات. وقد تتفق هذه الممارسة التعددية مع مختلف المنصات ونماذج الأعمال، كما يمكن أن تكون إيجابية لممارسة حُريّة التّعبير ومُكافحة المعلومات المُضلّلة، ذلك أن أي نموذج يتمتّع بطابع مركزي أكبر ويكون قابلاً للتنفيذ على الصعيد العالمي قد يأتي بنتائج معاكسة. وبين هذين النقيضَين، هناك مساحة متاحة للشركات لتشغيل ميزانها الأخلاقي الخاص بغية إقامة توازن بين ما تسمح بالتّعبير عنه وقرارات مراقبة المُحتوى التي تُتخذ لمُكافحة التّضليل والمحتويات الأخرى التي قد تعتبرها إشكالية من حيث سياساتها، و/أو التي تكون محفوفة بالمخاطر من الناحية القانونية في ولايات قضائية معينة.

وتشير مبادئ سانتا كلارا[248] إلى إطار محتمل للشفافية والمساءلة في مجال مراقبة المُحتوى. وقد طوّرت هذه المبادئ في أوائل العام 2018 مجموعة من الأكاديميين والمدافعين عن الحقوق الرقمية المعنيين بحُريّة التّعبير في مجال مراقبة المُحتوى على الانترنت في الولايات المتحدة. ويمكن أن تكون هذه المبادىء ذاتية التنظيم، غير أنها تستطيع أيضاً أن تسهم في السياسة التنظيمية. وهي تقترح معايير للإبلاغ عن الشفافية وآليات إشعار وطعن. ومن الأمثلة على إحدى التوصيات بشأن الطعون ضمان «توفّر مراجعة بشرية من جانب شخص أو فريق من الأشخاص لم يشاركوا في اتخاذ القرار الأولي». وتسعى المبادئ إلى التشجيع على اتباع مقاربة رفيعة المستوى قائمة على حقوق الإنسان إزاء مراقبة المُحتوى.

من جهته، ناصر المقرر الخاص للأمم المتحدة المعني بتعزيز وحماية الحق في حُريّة الرأي والتّعبير هذا النوع من المقاربات أيضاً. وقد نشر تقريراً عن مقاربة حقوق الإنسان في تنظيم مُحتوى المنصات

[248] https://santaclaraprinciples.org/، للاطلاع على أفكار حول ضمانات حُريّة التّعبير عند استخدام المراقبة المُميكنة للمُحتوى بغية التصدي للمعلومات المُضلّلة على الخط، انظر Marsden & Meyer (2019). يمكن أيضاً إيجاد الفقرات التالية من «مبادىء سانتا كلارا» ودراسة المقرر الخاص للأمم المتحدة في هذه الدراسة السابقة التي قُدّمت إلى البرلمان الأوروبي.

(Kaye, 2018). وعلى غرار الإعلان المشترك الصادر عن المقررين الخاصين للأمم المتحدة، ومنظمة الأمن والتعاون في أوروبا، ومنظمة الدول الأمريكية، واللجنة الأفريقية لحقوق الإنسان والشعوب بشأن حُريّة التّعبير و«الأخبار الزائفة» والتّضليل والدعاية (2017)[249]، يشير التقرير إلى ضرورة تحقيق التوازن عند تقييد حُريّة التّعبير (مع إيلاء الاعتبار الواجب للقانونية والضرورة والتناسب والمشروعية)، وحماية شركات التواصل عبر الإنترنت من المسؤولية عن مُحتوى الأطراف الثالثة. ويثير المقرر الخاص شواغل بشأن معايير المُحتوى التي تتعلق بالقواعد الغامضة، والكراهية، والمضايقة والإساءة، والسياق، ومتطلبات الاسم الحقيقي، والمعلومات المُضلّلة. ويحدّد التقرير سقفاً عالياً في هذا المجال وينصّ على مبادئ حقوق الإنسان ذات الصلة بمراقبة الشركات للمُحتوى (UN Special Rapporteur on Freedom of Opinion and Expression, 2018b, section IV, pars 44-63):

- حقوق الإنسان بشكل تلقائي، والقانونية والضرورة والتناسب وعدم التمييز عند التعامل مع مراقبة المُحتوى؛
- تفادي المخاطر التي تحدق بحقوق الإنسان والتخفيف من حدتها، والشفافية عند الاستجابة لطلبات الحكومة؛
- بذل العناية الواجبة، والانخراط والمدخلات العامة، وشفافية وضع القواعد عند وضعها وعند تطوير المنتجات؛
- المكننة والتقييم البشري، والإشعار والطعن، والإنصاف، واستقلالية المستخدم عند إنفاذ القواعد؛ والشفافية في اتخاذ القرارات.

على ذلك، أعرب المقرر الخاص عن قلقه إزاء «تفويض الوظائف التنظيمية إلى جهات فاعلة خاصة تفتقر إلى الأدوات الأساسية للمساءلة»، مشيراً إلى أن «عملياتها الحالية قد تتعارض مع معايير الإجراءات القانونية الواجبة، فيما تكون دوافعها اقتصادية أساساً» (الفقرة 17). وقد أشار التقرير أيضاً إلى أن «أشكال العمل الصريحة، مثل حجب المواقع الإلكترونية أو عمليات الإزالة المحددة، تنطوي على خطر التدخل الخطير في حُريّة التّعبير» (الفقرة 17)، وأن التدابير التكنولوجية التي تقيّد المُحتوى الإخباري «قد تهدد مصادر أخبارية مستقلة وبديلة أو مُحتوى تهكّمي. واتخذت السُّلطات الحكومية مواقف قد تعكس توقعات مبالغ فيها بشأن قدرة التكنولوجيا على حلّ هذه المشاكل بمفردها» (الفقرة 31).

جرت الإشارة إلى تحديات وفرص متعددة ترتبط بالاستجابات الهادفة إلى تنظيم المُحتوى رداً على المعلومات المُضلّلة التي تطلقها جهات فاعلة صحفية في القسمَين 6.1.2 و6.1.7 من هذا الفصل. وهي تركّز على تآكل وظائف الحارس التقليدية في عصر وسائل التواصل الاجتماعي، لا سيّما التحديات المتزامنة المتمثلة في بروز وتوزيع معلومات تهمّ المصلحة العامة وتكون ذات مصداقية ويمكن التحقّق منها وسط تسونامي من المعلومات المُضلّلة، والخطاب المسيء، والمُحتوى الموجه نحو الترفيه، علاوةً على الصحافة الحزبية المفرطة وذات النوعية الرديئة، وهي كلّها لعناصر تهدّد بإغراق مُحتوى حسن الصياغة والتنظيم يكافح المعلومات المُضلّلة. على ذلك، يطرح تنظيم المُحتوى الذي تضعه الجماهير على نطاق واسع على قنوات التواصل الاجتماعي المفتوحة وفي أقسام التعليقات المفتوحة – حيث تزدهر المعلومات المُضلّلة وخطاب الكراهية والإساءة – تحدياً كبيراً للغاية (Posetti et al., 2019b).

[249] انظر أيضاً القسمَين 5.1 و7.1 من هذا التقرير البحثي لمزيد من النقاش.

علاوةً على ذلك، تتوفّر تحديات أخلاقية ومهنية، مثل سوء تفسير مبدأ الموضوعية، حيث تُعامل المعادلة الكاذبة عن باب الخطأ كترياق للتحيز، الأمر الذي يؤدي إلى معاملة متساوية وغير نقدية لمصادر صادقة وغير صادقة. ويؤدي فقدان الثقة الناجم عن فشل النظام في وسائل الإعلام الإخبارية الى تقويض قدرة الصحافة المهنية على التصرّف كحصن ضد المعلومات المُضلّلة.

غير أن هذه التحديات بحدّ ذاتها تمثّل أيضاً فرصاً لناشري الأخبار والصحفيين لتمييز أنفسهم عن الآخرين بصفتهم منظمين مستقلين وأخلاقيين وناقدين لمعلومات ذات مصداقية وموثوقة وجديرة بالثقة تهمّ المصلحة العامة (Powell, 2020). كما أنها توفر فرصاً للابتكار في مجال إشراك الجمهور في المجتمعات المغلقة على وسائل التواصل الاجتماعي، مثل واتساب، للمساعدة في مُكافحة المعلومات المُضلّلة حيث يتم تداولها في غياب أي تدقيق وكشف واسع النطاق لزيفها (Posetti et al., 2019a).

10.1.6 توصيات بشأن الاستجابات الرامية إلى تنظيم المُحتوى

بالنظر إلى التحديات والفرص المحددة أعلاه وما يترتب على استجابات تنظيم المُحتوى من آثار كبيرة على حُريّة التّعبير، يمكن رفع التوصيات التالية في مجال السياسة العامة:

توصيات للدول الفردية:

- تعزيز الحاجة إلى إنشاء «مجالس لوسائل التواصل الاجتماعي» تكون مستقلة وتضمّ أصحاب مصلحة متعددين، على غرار مجالس الصحافة في قطاع الصحف، بالإضافة إلى وضع لوائح تنظيمية تتطلب الشفافية في طريقة تفسير شركات التواصل عبر الإنترنت لمعاييرها وتنفيذها، وتفتح المجال أمام الشكاوى التي تشمل القطاع برمّته، وتتيح للتعاون بين الشركات توفير سبل الإنصاف (UN Special Rapporteur on Freedom of Opinion and Expression, 2018b, pars 58, 59, 63, 72)[250].

توصيات للمنظمات الدولية:

- تشجيع شركات التواصل عبر الإنترنت على التأكد من أن استجابات تنظيم المُحتوى التي اعتمدتها شفافة وقابلة للقياس بشكل مناسب، وتدعم حقوق الإنسان، ويتم تنفيذها بشكل منصف (وتجنّب منح استثناءات لشخصيات سياسية نافذة مثلاً) على نطاق عالمي حقيقي.

توصيات لشركات التواصل عبر الإنترنت:

- تقديم تقارير مفصّلة ومتكررة عن الشفافية العامة، بما في ذلك معلومات محددة عن مشاهدة المعلومات المُضلّلة وانتشارها، وتعليق الحسابات التي تنشر معلومات مضللة، وتنفيذ عمليات الإزالة، وغير ذلك من الخطوات ضد المعلومات المُضلّلة، بما في ذلك حجب الربح المادي، نظراً إلى أن هذه الاستجابات تترك آثاراً هامة على حقوق الإنسان وحُريّة التّعبير.

[250] طُرحت فكرة مماثلة في Wardle (2017).

- إنشاء آليات مراجعة صلبة خارجية/تابعة لأطراف ثالثة من أجل مراقبة المُحتوى وضمان القدرة على الطعن في القرارات، بما في ذلك القرارات التي تقودها الآلة. ويشمل ذلك ضرورة مراجعة القرارات التي تقضي بعدم إزالة المُحتوى، بالإضافة إلى القرارات المتعلقة بحذفه.
- التأكّد من أن الاستجابات الرامية الى تنظيم المُحتوى تشجّع المستخدمين على الوصول إلى الصحافة من منظمات إخبارية مستقلة ومهنية أو غيرها من الجهات التي تنشر معلومات تخصّ المصلحة العامة وتكون نقدية وقائمة على الأدلة (مثل: الباحثين المستقلين ومنظمات المجتمع المدني الحسنة النية).
- تعزيز جهودها ضد هجمات التّضليل المدبّرة التي تُشن على الصحفيين من خلال استبعاد المستخدمين الذين يشاركون في هذه الاعتداءات على حُريّة الصحافة ويعيقون الجهود الرامية إلى التصدي للتضليل.
- اتخاذ خُطوات لضمان الدعم المناسب لمراقبي المُحتوى، بما في ذلك التدريب والأجور المتناسبة مع العمل المنجز وضمان الصحة النفسية.

توصيات لقطاع الإعلام:

- تسليط الضوء على المُحتوى المضاد للمعلومات المُضلّلة (كالمُحتوى الذي يساعد على تثقيف الجماهير حول مخاطر المعلومات المُضلّلة، ويساعد على تجهيزهم لمقاومته ومواجهته حيث يجدونه، ويبرز آليات كشف مهمة مثل دحض الأساطير بشأن جائحة فيروس كورونا).
- اختبار وسائل مبتكرة تتيح للجمهور تنظيم المُحتوى والانخراط، خاصة داخل تطبيقات مغلقة تزدهر فيها المعلومات المُضلّلة.
- دعوة شركات التواصل عبر الإنترنت وهيئات الحوكمة ذات الصلة إلى القيام بتدخلات تنظّم المُحتوى وتتصدى للمعلومات المُضلّلة لمراعاة الأطر الدولية لحقوق الإنسان، والحرص على امتثال القيود المفروضة في حالات الطوارئ (مثل جائحة فيروس كورونا) لشروط المعايير الدولية بشأن تقييد الحقوق.
- تأمين مراقبة نقدية لجهود تنظيم المُحتوى التي تبذلها شركات التواصل عبر الإنترنت لتعزيز الشفافية والمساءلة.

ملاحظة: ترد توصيات أخرى خاصة بتنظيم مُحتوى الإعلانات وحجب الربح المالي في القسم 6.3.

2.6 الاستجابات التقنية/الخوارزمية

المؤلفون: سام غريغوري، وكالينا بونتشيفا، وتريشا ماير، ودينيس تيسو

يستعرض هذا القسم أحدث الخوارزميات والتكنولوجيا من أجل الكشف (شبه) المُميكن عن المعلومات المُضلّلة على الإنترنت وفائدتها العمليّة عبر دورة حياة حملات التّضليل، بما في ذلك تحليل المُحتوى ومصداقية المصدر، وانتشار الشبكات، وقياس التأثير على معتقدات المواطنين وأفعالهم، وأساليب كشف الزيف. وتهدف هذه التدابير التقنية، بدرجة أكبر أو أقل، إلى تعزيز بروتوكولات تنظيم المُحتوى التي تعتمدها الشركات أو غيرها من بروتوكولات السياسة العامة أو حتى تنفيذها. ويتيح استخدام التدابير التقنية خارج إطار الشركات، من قبل المجتمع المدني و/أو الأكاديميين وغيرهم من الجهات الفاعلة، تقييم قضايا مثل وجود معلومات مضللة (وأنواع أخرى من المُحتوى) وتدفّقها. ويشير طابع الاستجابات التقنية/الخوارزمية «التالي للتنفيذ هذا» إلى أن التحديات أو الفرص المتاحة لحُريّة التّعبير والناشئة عن تطبيق التكنولوجيا قد تنتج عن السياسات الرسمية أو غير الرسمية في المرحلة التمهيدية المطروحة. ولابدّ من الإشارة إلى أن عدم إدراج مبادئ حُريّة التّعبير في مرحلة تصميم الاستجابة التقنية يمكن أن يحدّ من فعالية الاستجابة أو يتسبّب بآثار سلبية غير مقصودة. وفي الوقت نفسه، قد تبرز مشاكل أيضاً من منظور حُريّة التّعبير عندما يرتبط منطق التصميم التكنولوجي ارتباطاً مباشراً ضئيلاً بمنطق السياسات/الغرض، فيعمل بشكل مستقل عن هذا الاتجاه.

يمكن أن تقوم المنصات الاجتماعية ومحركات البحث في حدّ ذاتها بتنفيذ هذه الاستجابات التقنية/الخوارزمية، ولكن يمكن أن تضطلع بذلك أيضاً أدوات خارجية أو تابعة لأطراف ثالثة (مثل البرامج المساعِدة للمتصفّح) أو أساليب تجريبية ناجمة عن البحث الأكاديمي. وتشمل التكنولوجيا التي تُناقش في هذا القسم من الدراسة قواعد بيانات الهاش التقنية hash databases، والترتيب المُميكن، ومصفاة التحميل، من بين أمور أخرى. وتخضع أيضاً للاستعراض التكنولوجيا الناشئة حديثاً والدراية في مجال تحليل المُحتوى المغلوط الذي تم إضفاؤه بصورة تلقائية (أي الصور الزائفة الفجّة deepfakes أو المادة الإعلامية المركّبة) عبر الصوت والنصّ والصور ومقاطع الفيديو. ويتناول هذا القسم أيضاً الوسائل التكنولوجية لتحديد «السلوك المُفتعل المُعد له» و«الجهات الفاعلة المُفتعلة» والتصرف على أساسها، علماً أنها مقاربة تختلف عن تحديد المُحتوى وتُكمّله. والمقاربة هذه تقضي بتحديد الأنماط التي تميل إلى الارتباط بحملات التّضليل من الناحية التكنولوجية.

بالإضافة إلى ذلك، يتم تحليل نقاط القوة والضعف والثغرات في مجموعة من المقاربات الأخرى القابلة للتحقُّق من المُحتوى وإجراء تحليل جنائي إعلامي. ويتمثل أحد التحديات الهامة بشكل خاص في تحقيق التوازن بين مراعاة شفافية الخوارزمية (مثل: الحرص على إمكانية التحقُّق من الخَيارات الخوارزمية، وفهم التحيزات الضمنية والصريحة) وخطر إضعاف فعالية الخوارزمية، الأمر الذي من شأنه أن يتيح للجهات الفاعلة في مجال التّضليل استغلال نقاط الضعف وتطوير استراتيجيات للإفلات. وتُطرح في هذا الصدد أيضاً مسألة أخرى هي إمكانية الوصول إلى الأدوات التي تعتمد على المقاربات الخوارزمية.

1.2.6 مَن وما هي أهداف الاستجابات التقنية والخوارزمية؟

ترصد الاستجابات التقنية والخوارزمية نطاق المعلومات المُضلّلة وطبيعتها، وتستخدم التقنية الآلية لدعم عملية صنع القرار داخل منصات التواصل على الإنترنت ولصالح أطراف ثالثة. وهي توفر مقاربات لتقييم مصداقية عناصر المُحتوى ومصادره، والسلامة الإعلامية لأشكال جديدة من المواد الإعلامية المركّبة، فضلاً عن رصد تدفق المعلومات والنشاط الحسابي، كاستخدام روبوتات الويب.

2.2.6 مَن الذي تسعى الاستجابات التقنية والخوارزمية إلى مساعدته؟

تدعم الاستجابات التقنية في المقام الأول الكثير من أصحاب المصلحة، كمنصات التواصل عبر الإنترنت، فضلاً عن وسائل الإعلام، والجهات التي تتحقّق من الوقائع، والمحققين. وتتيح أدوات مشاركة الصور والفيديو ومُحرّكات البحث والرسائل لمنصات الإنترنت نفسها أن تجري عمليات شبه آلية للكشف عن الرسائل والوكلاء وكيفية انتشار المُحتوى، فضلاً عن تأمين المعلومات لأطراف أخرى (كالجهات الخارجية التي تتحقّق من الوقائع). وتتوفّر مجموعة ذات صلة من الأدوات تدعم الصحفيين ووسائل الإعلام والجهات التي تتحقّق من الوقائع والمحققين الذين يُجرون تحقيقات محددة أو يوثّقون نطاق التّضليل على المنصات.

تُعدّ معظم الأدوات الآلية المُستخدمة في الكشف عن المعلومات المُضلّلة مناسبة لتأمين المدخلات لعملية صنع القرار البشرية -إما على مستوى عنصر المُحتوى أو تقييم نمط سلوك الجهات الفاعلة. فعلى مستوى المُحتوى، توفّر معلومات تتيح التحليل البشري للمصدر والتلاعب. أما على مستوى الجهات الفاعلة، فهي تقدّم معلومات عن نشاط روبوتات الويب عبر الإنترنت المحتملين وعن النشاط المشبوه المتصل بالشبكة.

ويتمثّل الافتراض وراء المقاربات التقنية والخوارزمية في أنها قادرة على تقليص وجود المعلومات المُضلّلة ومشاركتها والحدّ من الحوافز التي تحرّك الجهات الفاعلة في حقل التّضليل. وتقضي نظرية التغيير الحالية بأنه على ضوء الكمّ الهائل من المعلومات وضرورة الكشف عن الحملات المُعد لها أو عمليات التلاعب الفردية التي لا يستطيع الإنسان التعرّف إليها بسهولة، باستطاعة الأدوات التقنية الآلية أن تساعد في فرز عملية اتخاذ القرارات، والحد من الاهتمام المزدوج، وتسريع القرارات الفردية وعمليات توفير المعلومات. ومع ذلك، يبدو أن الهدف على المدى الطويل هو وضع مقاربات خوارزمية ومقاربات يسيّرها التعلّم الآلي تتّسم بفعالية أكبر وتقلل من ضرورة إجراء مراقبة وتحليل العنصر البشري للمُحتوى (وتوفّر الموارد البشرية والمالية المطلوبة)، فتتيح بالتالي إجراء تنظيم آلي أكبر للمُحتوى من دون العودة إلى مراقبين بشر، كما هو الحال مع المقاربات المعتمدة إزاء التطرف العنيف عبر الإنترنت.

من المرجح أن يأتي الانتقال إلى مراقبة آلية إضافية للمُحتوى -وهو انتقال فرضته جائحة فيروس كورونا- وضرورة العمل مع قوى عاملة بشرية أقلّ تعمل عن بُعد - كما ذكر فيس بوك[251] وتويتر[252]

....

251 https://about.fb.com/news/2020/03/coronavirus/#content-review

252 https://blog.twitter.com/en_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html

ويوتيوب[253]– بنظرة متعمّقة حول الموضوع على المدى القصير (بشرط أن تتحلّى الشركات ببعض الشفافية بشأن ما يحدث خلال هذه التجربة القسرية). وقد أشار فيس بوك، في تدوينة حول هذه المسألة، إلى أنه مع «تقليص اليد العاملة وتأديتها عملها عن بُعد، سوف نعتمد اليوم اعتماداً أكبر على أنظمتنا الآلية للكشف عن المُحتوى المخالف وإزالته وتعطيل الحسابات. ونتيجةً لذلك، نتوقع ارتكاب المزيد من الأخطاء، وسوف تستغرق المراجعات وقتاً أطول من المعتاد، ولكننا سنستمرّ في رصد أداء أنظمتنا وإجراء تعديلات». يعكس هذا البيان فهماً بأنّ النظم الآلية ليست في الوقت الحالي بديلاً عن الرقابة البشرية، وأنها تتطلّب تصويبات وطعون قوية «كما أشار إليه مقرر الأمم المتحدة الخاص المعني بتعزيز وحماية الحق في حُريّة الرأي والتّعبير» (Kaye, 2018).

3.2.6 ما هي المخرجات التي تنشرها الاستجابات التقنية والخوارزمية؟

يمكن القول، بشكل عام، إنه على عكس الأنظمة الآلية التي بُنيت للكشف عن صور استغلال الأطفال أو المُحتوى المتطرف العنيف والتي تحذف المُحتوى إلى حد كبير دون رقابة بشرية على كل قرار، توفّر أنظمة الكشف عن المعلومات المُضلّلة على نطاق واسع معلومات لإطلاق عمليات صنع القرار البشرية اللاحقة داخل شركات الإنترنت بشأن الاستجابة لحملات التّضليل أو تصنيف مُحتوى أو حسابات محددة أو خفض ترتيبها أو إزالتها.

على الرغم من أنّ معظم منصات الإنترنت الكبرى تكتب في الوقت الحالي تقارير شفافية بشأن مستويات إزالة المُحتوى أو الحسابات، فضلاً عن تقارير تحقيقية بشأن النتائج المسجّلة في مواجهة حملات تضليل معينة (انظر القسم 4.2 لمزيد من التفاصيل)، فإنّ هذه التقارير لا تتضمّن شفافية متعمقة بشأن تبعات استخدام الخوارزميات والتعلم الآلي وغير ذلك من أشكال صنع القرار الآلي على حقوق الإنسان. كما أن هذه التقارير لا تشرح المعايير التي تُعتبر هذه الأساليب على أساسها تدخلات فعالة. ويشمل نطاق الإفصاح إجمالاً أرقاماً عامة حول استخدام النظم الآلية وتنفيذها –على سبيل المثال في تقرير صدر مؤخراً[254]، يشير فيس بوك إلى قدرته على تحديد 99٪ من الحسابات الزائفة بشكل استباقي (أي آلياً دون تدخل بشري). وتعتبر المنصات أن هذا هو المستوى المناسب من الشفافية، نظراً إلى الطابع الخلافي لمراقبة المُحتوى وللأسلوب الذي ستعتمده «الجهات الفاعلة السيئة» لمحاولة استغلال فهم الخوارزميات المُستخدمة لمراقبة المُحتوى.

نظرت دراسة أجرتها مبادرة Ranking Digital Rights في مسألة الشفافية في محركات التوصيات (Ranking Digital Rights, 2020). واستعرضت خمس شركات إنترنت كُبرى، بما في ذلك آبل (iOS) وغوغل (بحث، يوتيوب، أندرويد)، وفيس بوك (فيس بوك)، ومايكروسوفت (بينغ Bing، وان درايف OneDrive) وتويتر، ووجدت ثغرات في الحوكمة وضعف في مجال العناية الواجبة بحقوق الإنسان. ويشير التقرير إلى أن «أياً من المنصات الخمس التي تتّخذ من الولايات المتحدة الأميركية مقراً لها والتي تم تقييمها؛ قدّم التزامات عامة صريحة بحماية حقوق الإنسان أثناء تطوير النظم الخوارزمية واستخدامها» وأن «الشركات التي تدير منصات عالمية رئيسية لا تقدم أدلة تثبت أنها تجري تقييمات للمخاطر تمكّنها

253 https://youtube-creators.googleblog.com/2020/03/protecting-our-extended-workforce-and.html

254 https://transparency.facebook.com/community-standards-enforcement

من فهم الأضرار التي تحدق بحقوق الإنسان وتنجم عن استخدامها للنظم الخوارزمية ونماذج الأعمال الهادفة المستندة إلى الإعلانات على مستخدمي الإنترنت ومن التخفيف من حدّتها». وحدها شركة أمريكية واحدة (مايكروسوفت) كشفت أنها تجري تقييمات لتأثير تطوير أنظمة خوارزمية واستخدامها. ولم تكشف أي من الشركات الثماني في الدراسة عما إذا كانت تجري تقييمات للمخاطر تتناول كيفية تأثير سياساتها وممارساتها الإعلانية الهادفة على حُريّة المستخدمين في التّعبير وحقّ الوصول إلى المعلومات، أو حقهم في احترام الخصوصية أو عدم التمييز.

تختلف أنظمة الأطراف الثالثة القابلة لاستكمال التدقيق في المُحتوى أو تحديد أشكال جديدة من الوسائط الاصطناعية من حيث درجة تطور مخرجاتها. وهناك عدد من أدوات الأطراف الثالثة، مثل INVID وAssembler، التي تدمج مجموعة من الأدوات مفتوحة المصدر في لوحات المعلومات لمساعدة الصحفيين المحترفين والمحققين.

1.3.2.6 المُقاربات داخل الشركات التي تتناول وسائل التواصل الاجتماعي، وتبادل الفيديو، ومحركات البحث، والرسائل من أجل الكشف (شبه) الآلي عن حملات التّضليل على الإنترنت، بما في ذلك الأدوات الآلية للكشف وقواعد بيانات الهاش التقنية databases Hash ومصفاة التحميل

تقوم شركات ومنصات الإنترنت بنشر مجموعة من نماذج الكشف الآلي عن أنواع المُحتوى على خدماتها، وهي تشمل أدوات لتتبُّع الانتشار العضوي والاصطناعي للمعلومات، ولتحديد المُحتوى الذي يفي بمعايير خفض الترتيب أو التصنيف أو الحذف.

أدوات آلية للكشف عن وإدارة سلوك يعتمد التّضليل

يمكن استخدام أسلوب التعرّف الآلي إلى المُحتوى لإصدار قرارات آلية وتنفيذها أو لمساعدة الإنسان في اتخاذ قرارات بشأن مراقبة المُحتوى أو تحديد الأنماط. وكما أُشير إليه في تقرير الاتحاد الأوروبي المعنون «تنظيم المعلومات المُضلّلة بواسطة الذكاء الاصطناعي» (Marsden & Meyer, 2019)، «من بين تقنيات التعلُّم الآلي التي تتقدم باتجاه الذكاء الاصطناعي، تُعدّ تقنيات التعرف الآلي إلى المُحتوى برامج تحليل نصية وسمعية-بصرية يتم تدريبها خوارزمياً لتحديد حسابات روبوتات الويب المحتملة ومواد التّضليل المحتملة الغريبة». ويعترف التقرير بأن مراقبة المُحتوى على نطاق أوسع تتطلّب تقنيات التعرف الآلي إلى المُحتوى التي تكمل المراقبة البشرية للمُحتوى (التحرير)، ولكنه يذكر أن استخدام هذه التقنيات للكشف عن المعلومات المُضلّلة عُرضة للسلبيات/الإيجابيات الكاذبة بسبب صعوبة تحليل معان متعددة ومعقدة ولربما متضاربة تنشأ عن النص. وإذا كانت تقنيات التعرّف الآلي إلى المُحتوى هذه غير مناسبة لمعالجة اللُّغة الطبيعية وحتى المواد السمعية- البصرية، بما في ذلك «الصور الزائفة الفجّة» deep fakes (أي التمثيل الاحتيالي للأفراد في الفيديو)، فإنها لاقت نجاحاً أكبر في التعرّف إلى حسابات «روبوتات الويب»، وفقاً للتقرير.

على الرغم من أن خوارزميات الكشف الفعلية التي تُستخدم في المنصات للكشف عن مُحتوى أو سلوك مُفتعل ليست متوفرة للتدقيق العام، فقد دمج تويتر مقاربات كشف لمعرفة ما إذا كان حساب ما يستخدم صورة رمزية avatar مسروقة، أو نص مسروق أو منسوخ من ملف تعريف، أو موقع مضلل لملف التعريف (Harvey & Roth, 2018). من جهته، يحتوي فيس بوك على عدد أقل من حسابات روبوتات الويب

الآلية ولكن يتعيّن عليه تحديد المزيد من «الحسابات الدُّمى» sock puppets (حسابات كاذبة متعددة مع إنسان حقيقي وراءها) وحسابات انتحال الشخصية بدلاً من ذلك. ويُعتبر تحديد هذه الحسابات آلياً أصعب بكثير من العثور على روبوتات الويب (وقد يكون مستحيلاً في بعض الأحيان)، بسبب الطابع الأكثر أصالة للسُّلوك الذي يُحرّكه الإنسان (Weedon et al., 2017). إلى جانب ذلك، تستخدم أحدث الأبحاث بشأن طرق الكشف عن روبوتات الويب ميزات السلوك الاجتماعي بالدرجة الأولى –مثل تواتر التغريدات، واستخدام الوسم، ومتابعة عدد كبير من الحسابات في حين لا يتبع هكذا حسابات إلا عدد قليل فقط (Varol et al., 2017; Woolley & Howard, 2016; Cresci et al., 2016). وتتوفّر أيضاً مقاربات تكتشف روبوتات الويب على أساس الارتباط الكبير في الأنشطة بينها (Chavoshi et al., 2017).

يستخدم موقع ويكيبيديا، الذي بُني على أساس إسهامات معرفية يسهم بها المستخدمون، روبوتات الويب (أي وكلاء آليين)[255] «للقيام بدوريات» على صفحاته وتحديد السلوك الذي يُعتبر أنه يهدف عمداً «إلى عرقلة الغرض من المشروع أو إفشاله، علماً أن هذا الغرض هو إنشاء موسوعة حرة متوفرة بمجموعة متنوعة من اللغات، تعرض مجموع المعارف البشرية»[256]. وقد قدم مجتمع ويكيبيديا سلسلة من المقترحات حول كيفية إنشاء روبوتات الويب للتعامل مع «الحسابات الدُّمى» المُستخدمة لإجراء تعديلات، كما قد يحدث في سياق حملة تضليل مُعد لها، ولكن يبدو أن هذه المقترحات لم تُنفّذ بعد.

أدوات آلية للتعرّف إلى المُحتوى وإزالته، بما في ذلك قواعد بيانات الهاش التقنية hash databases ومصفات التحميل:

تُستخدم الأدوات الآلية لإزالة المُحتوى، مثل قواعد بيانات الهاش التقنية hash databases والبصمات، بشكل أساسي في سياق صور استغلال الأطفال والصور المحمية بحقوق الطبع والنشر (مثل: YouTube Content ID) والمُحتوى المتطرف العنيف، لا سيما في سياق الولايات القانونية، لتحديد هذا المُحتوى وإزالته. وتتيح قاعدة بيانات الهاش التقنية للمنصات تحديد النسخ المكرّرة أو شبه المكررة، استناداً إلى أوجه تطابق مع عناصر مُحتوى موجودة في قاعدة بيانات.

وتقنية الهاشينغ Hashing هي تقنية تنطوي على تطبيق خوارزمية حسابية لإنتاج قيمة فريدة تمثّل أي مجموعة من الوحدات متناهية الصغر «bits»، كصورة أو فيديو. وتتوفر مجموعة متنوعة من مقاربات الهاشينغ Hashing، بما في ذلك تجزئة كل إطار من شريط فيديو أو فترات منتظمة من الإطارات، أو تجزئة أقسام فرعية من صورة ما. ويمكن أن تساعد تقنيات الهاشينغ هذه في الكشف عن التلاعب –مثل ما إذا كانت صورة ما قُصّت– وفي تحديد مجموعات فرعية من اللقطات التي تم تحريرها والتحقّق منها. وتقوم أدوات مثل تكنولوجيا PhotoDNA المستخدمة عبر الشركات لدراسة صور استغلال الأطفال باحتساب قيم الهاش hash ، استناداً إلى المُحتوى المرئي للصورة (عن طريق تحويل الصورة إلى أبيض وأسود، وتغيير حجمها، وتقطيعها على شكل شبكة، والنظر في تدرجات الكثافة أو الحواف). وهي بالتالي تؤدي دوراً أفضل في اكتشاف مواد إعلامية خضعت لتعديلات، وليس مجرد النسخ الدقيقة.

255 https://en.wikipedia.org/wiki/Wikipedia:Bots

256 https://en.wikipedia.org/wiki/Wikipedia:Vandalism

حتى وقت قريب، لم تكن هناك آلية رسمية مُعد لها بين شركات ومنصات الإنترنت لرصد المعلومات المُضلّلة أو استخدام مقاربة مشتركة تتعلّق بالهاش hash أو البصمات في هذا المجال، على عكس حال التطرف العنيف؛ حيث يجري التنسيق من خلال كيانات مثل «المنتدى العالمي للإنترنت لمُكافحة الإرهاب» الذي تتمثّل فيه معظم الشركات الكبرى. وفي آذار/مارس 2020، أعلن كل من فيس بوك وغوغل ولينكدن ومايكروسوفت وريديت وتويتر ويوتيوب بشكل مشترك أنها تعمل معاً عن كثب للاستجابة لجائحة لفيروس كورونا و«لمُكافحة الاحتيال والمعلومات المغلوطة بشأن الفيروس بشكل مشترك». ولم يتّضح ما إذا كان هذا التعاون يتضمّن مقاربة هاش hash مشتركة (Facebook, 2020a). ولم يتضح أيضاً كيف تستطيع مقاربة من هذا النوع، إذا تم توسيعها لتتجاوز المعلومات المغلوطة والمُضلّلة حول فيروس كورونا، أن تربط بين السياسات/معايير المجتمع المختلفة للشركات في هذا المجال (على سبيل المثال، لفائدة كيفية التعامل مع الإعلانات السياسية التي تحتوي على أكاذيب، أو كيفية إدارة مواد إعلامية تمّ التلاعب بها) أو مجموعة الأساليب التي يتحول من خلالها المُحتوى المغلوط/المضلل كلما قام المستخدمون بتحريره وتغييره. وبالمثل، يبدو أن التنسيق بين وسائل الإعلام ومنصات الإنترنت الرئيسية في إطار مبادرة Trusted News Initiative لا يتضمن أي مقاربة هاش hash أو مُطابقة.

غالباً ما تُستخدم مصفات التحميل بالتزامن مع تقنية الهاشينغ hashing والبصمات، فتقوم بتقييم المُحتوى في نقطة التحميل لمنع المشاركة، ويتم استخدامها أقلّ في سياق المعلومات المُضلّلة. وتبرز مخاوف كبيرة بشأن حُريّة التّعبير متى تم استخدام مقاربتَي الهاشينغ hashing والبصمات، لا سيما بالاقتران مع مصفاة التحميل. وتشمل هذه المخاوف الشفافية حول كيفية إضافة صورة معينة إلى قاعدة بيانات الهاش أو البصمات، فضلاً عن مخاوف حول كيفية النظر في السياق المحيط بصورة ما (كما هو الحال مع المُحتوى الحقيقي الموزّع بطرق تديم التّضليل، مع تعليق أو وصف غير دقيق مثلاً على وسائل التواصل الاجتماعي). وكما لاحظ باحثان، «إنّ التقنيات الآلية محدودة في دقتها، خاصةً في مجال التّعبير عندما تكون الإشارات الثقافية أو السياقية ضرورية. وبالفعل، يُسهّل تحديد عدم مشروعية المُحتوى الإرهابي أو المُحتوى المسيء للأطفال مقارنةً بحدود الخطاب السياسي أو أصالة الأعمال المشتقة (الخاضعة لحقوق الطبع والنشر). وينبغي ألا نُلقي بهذا الإجراء الصعب الذي يتطلّب الحكم على المعلومات المُضلّلة على عاتق وسائط الإنترنت» (Marsden & Meyer, 2019).

تعكس المخاوف المتعلقة بمصفات التحميل (والسبب وراء عدم إمكانية استخدامها حالياً لرصد المعلومات المُضلّلة) حقيقة أن برمجية رصد التحميل لا تستطيع أن تميّز القصد، مثل التهكُّم والتقليد الساخر اللذين قد يُعيدان توظيف المُحتوى الموجود (للاطلاع على أمثلة أخرى، انظر Reda, 2017). والأمر الذي يعزّز المخاوف هو غياب الشفافية بشأن المُحتوى الذي تشير إليه هذه المصفات. ويجري استخدام مصفات التحميل الحديثة في مجالات أخرى من مراقبة المُحتوى -لا سيما في إطار إنفاذ حقوق التأليف والنشر وبصورة متزايدة في مجال مُكافحة الإرهاب والتطرف العنيف- ولكن ليس في حقل المعلومات المُضلّلة.

أدوات متوفّرة لوسائل الإعلام والمجتمع المدني للانخراط مع أنظمة المنصات

تستثمر بعض شركات الإنترنت أيضاً في أدوات تمكّن أطرافاً ثالثة من المساهمة بشكل أفضل في تحديد المُحتوى أو التحقُّق من وقائعه. وكما نوقش في القسم 4.1، يدعم فيس بوك شبكة من الجهات الخارجية التي تتحقّق من الوقائع، والتي يتم تزويدها بقائمة بالقصص التي وسمها المستخدمون والتي

حددتها فرق مراجعة المُحتوى الداخلية في فيس بوك. بالإضافة إلى ذلك، تستطيع هذه الجهات التي تتحقق من الوقائع إضافة قصص أُخرى حدّدتها بنفسها للتحقق من المصداقية (على الرغم من أنّ هذا العمل ليس من تلقاء نفسه، فهي ستتقاضى أجراً مقابله). ويقول فيس بوك أنه يقلّل بعد ذلك بنسبة 80٪ من إمكانية رؤية القصص التي تعتبرها الجهات التي تتحقّق من الوقائع كاذبة (DCMS HC 363, 2018b) وأنه يحدّ أيضاً من نطاق عمل المجموعات التي تشارك معلومات مضللة بشكل متكرر (Rosen & Lyons, 2019).

Claim Review[257] هو مخطط ترميز لصفحات الويب طوّره غوغل و Duke Reporters' Lab لإتاحة المجال أمام وسم أسهل لقصص تضمّ معلومات عن واقع أساسي تم فحصه، ومعلومات عن الطرف الذي أدلى بالمعلومات، وحكم بشأن دقّتها. ويجري الآن تطوير نسخة من هذه المقاربة اسمها MediaReview لتمكين الجهات التي تتحقّق من الوقائع من تأمين وسم أفضل لأشرطة الفيديو والصور المغلوطة (Benton, 2020).

كما نوقش في القسم 7.3 حول استجابات التمكين وتصنيف المصداقية، تنظر مجموعة من الشركات في إمكانية تطوير أدوات للمصادقة على المُحتوى وإسناده وتتبّع مصدره، وتطوير هيكلية البحث والتأكد من أصله. ومن الأمثلة على ذلك، مبادرة Content Authenticity Initiative التي وضعها كل من أدوبي وتويتر ونيويورك تايمز بهدف إنشاء «إطار إسناد مفتوح وقابل للتوسُّع... تستطيع أي شركة أن تنفذه ضمن منتجاتها وخدماتها» (Adobe, 2019).

2.3.2.6 أدوات متوفّرة لوسائل الإعلام والمجتمع المدني لفهم وكلاء التّضليل ووسطائه وأهدافه، ولتعزيز عمليات تقييم التلاعب والتحقُّق من الوقائع

كشف الأطراف الثالثة عن وكلاء التّضليل والسلوك والشبكات ذات الصلة

يقضي أحد الجوانب الرئيسية لتحليل المعلومات المُضلّلة بتحليل الوكلاء الذين يقفون وراء حملات التّضليل، والوكلاء الرئيسيين الآخرين المعنيين، والروابط الشبكية الصريحة أو الضمنية بينهم. ويتمثّل أحد الجوانب الأساسية لذلك في معرفة جدارة وكلاء التّضليل هؤلاء بالثقة ومصداقيتهم. ويصف بعض الباحثين هذا الأمر بأنه «تحقق من المصدر»، ويعتبرونه بالغ الأهمية خاصةً من حيث المساعدة التي توفّرها الأدوات والمقاربات الآلية، على الرغم من التغاضي عنه في الوقت الحالي (Wardle & Derakhshan, 2017). وقد اقترحت أبحاث صحفية عدة مقاييس لتقييم جودة الأخبار والمواد الإعلامية عبر الإنترنت، مثل التحيز الحزبي، والتحيز الهيكلي، والتحيز الموضعي، وشفافية المصدر (Lacy & Rosenteil, 2015). غير أنه لا توجد في الوقت الحالي طرق آلية لحسابها. وقد نال التعرّف الآلي إلى التحيز الإعلامي في المقالات الإخبارية الاهتمام في دراسة استقصائية أُجريت مؤخراً (Hamborg, Donnay & Gipp, 2018). وتكمن المؤشرات المتعلقة بجدارة المصادر القائمة على المُحتوى بالثقة؛ المؤشرات التي باتت حالياً مفهومة بشكل أفضل والناجمة عن أبحاث الكشف عن روبوتات الويب. ويُناقش عدد من هذه المبادرات القائمة على تقييم مصداقية الجهات الفاعلة، مثل المؤشر العالمي للتضليل Global Disinformation Index[258] في أقسام أخرى (لا سيّما القسم 7.3).

257 https://schema.org/ClaimReview

258 https://disinformationindex.org/

لا يتصرّف وكلاء التّضليل إجمالاً بشكل مستقل، على الرغم من صعوبة إثبات هذه المسألة في بعض الأحيان. فللإيحاء بأن عدداً كبيراً من المصادر المستقلة تقوم بالإبلاغ بطرق مختلفة عن «الوقائع» نفسها، تقوم بعض مواقع التّضليل و/أو الحسابات الدمى sock puppet بإعادة استخدام مُحتوى مواقع أخرى وإعادة نشره، في ممارسة تُعرف باسم غسيل المعلومات (Starbird, 2017). ويفتقر الصحفيون حالياً إلى أدوات سهلة الاستخدام تُظهر المواقع الإعلامية البديلة أو حسابات الشبكات الاجتماعية التي أعادت استخدام مُحتوى من موقع آخر. وهذا أمر مهم، لأن وسائل الإعلام الحزبية المفرطة والحسابات الدُّمى sock puppet تقوم بإعادة تنسيق المُحتوى و/أو إعادة نشره في محاولة لنيل المصداقية والقبول من خلال الطابع الأليف. وقد ركّزت الأبحاث حتى الآن في المقام الأول على دراسة إعادة التغريد وأنماط الذّكر في أوساط هكذا جهات كاذبة تعمل على تضخيمها، مثلاً في الانتخابات الرئاسية الأمريكية لعام 2016 (Faris et al., 2017)، غير أننا ما زلنا نفتقر إلى التكنولوجيا التي تُتيح إجراء تحليل أكثر عمقاً بكثير.

التحليل الآلي لرسالة/مُحتوى من جانب أطراف ثالثة

من بين الشركات الناشئة التي تعمل على مقاربات الكشف اعتماداً على الذكاء الاصطناعي لتقييم جودة المُحتوى أو مؤشرات تدلّ على أن عنصر مُحتوى ما مُلفّق، يمكن ذكر Factmata [259] وiAdverifa [260]. بالإضافة إلى ذلك، حدّدت ائتلافات مثل Credibility Coalition مؤشرات قائمة على المُحتوى لمصداقية الرسائل كنقطة انطلاق للتمديدات المحتملة لمعايير مخطط الويب الحالية. وتتضمن مؤشرات المُحتوى الرئيسية المتعلقة بالتّضليل عناوين مثيرة للانتباه clickbait وبعض المُغالطات المنطقية. وتتداخل هذه المقاربات مع مسائل تُناقش في القسم 7.3 وهي ليست حتى الآن مُنشأة آلياً.

أدوات الأطراف الثالثة للكشف عن روبوتات الويب، والتضخيم الحسابي، والحسابات الزائفة، أو لإنشاء معلومات موثوقة حول المُحتوى المجمّع أو المستند إلى التعلّم الآلي:

في حين تُبقي شركات الإنترنت الكبرى على الغموض لفائدة عملياتها القابلة للكشف عن روبوتات الويب، يتوفّر عدد من الأدوات التي طوّرتها الشركات والمجتمع المدني والأوساط الأكاديمية.

تتوفّر خدمة واسعة الانتشار تابعة لتويتر للكشف عن روبوتات الويب، وهي Botometer [261] (BotOrNot سابقاً)، التي تؤمنها جامعة إنديانا مجاناً. يستطيع المستخدمون التحقُّق من معدل احتمال وجود روبوتات الويب في حساب تويتر معين، استناداً إلى معلومات من ملف التعريف عن المستخدم وأصدقائه ومتابعيه. ويخضع الاستخدام لمصادقة تويتر ولمعدّل مقيّد من عدد الطلبات التي يمكن تقديمها إلى واجهة برمجة التطبيقات API الخاصة بتويتر. وبشكل عام، بما أن الأساليب القائمة على الأبحاث تستخدم البيانات التي تم الكشف عنها علناً بشأن حسابات تويتر فقط، تتوفر مخاوف حول دقتها، نظراً إلى أنّ منظّمي المُحتوى البشر يمكن أن يناضلوا في الكثير من الأحيان للتعرف إلى روبوتات الويب من ملفات التعريف العامة التي تخصّ تويتر وحده، وأن يقترفوا الأخطاء في مجال سوء الإسناد. ومن

....

[259] https://factmata.com/

[260] https://adverifai.com/

[261] https://botometer.iuni.iu.edu/#!/

المتوقّع أن تصبح الأمور أكثر صعوبةً مع بدء ظهور روبوتات ويب أكثر تطوراً. وتستعرض أعمال أُجريت مؤخراً؛ التحديات التي تطرحها أجهزة الكشف الآلية عن روبوتات الويب مع مرور الوقت، مشيرةً إلى مشكلات تباين من حيث الإيجابيات والسلبيات الخاطئة، لا سيما خارج إطار اللُّغة الإنجليزية، الأمر الذي أدى إلى دراسات «تعتبر بلا علم عدداً كبيراً من المستخدمين البشر كروبوتات ويب والعكس» .(Rauchfleisch & Kaiser, 2020)

خلال الانتخابات في البرازيل، قام فريق من الباحثين في UFMG بتنفيذ 'Bot o Humano' مستخدماً الوصول إلى واجهة برمجة التطبيقات API لتويتر من أجل توفير خدمة كشف[262] تُركّز على كيفية قيادة روبوتات الويب الموضوعات الأكثر رواجاً. علاوة على ذلك، قدم الباحثون خدمات ذات صلة لرصد مجموعات سياسية عامة على واتساب[263] (Melo & Messias et al., 2019)، وقد أصبحت متاحة في وقت لاحق للاستخدام في الهند وإندونيسيا ولرصد الإعلانات على فيس بوك (Silva & Oliveira et al., 2020). ويوفّر مقدمو الخدمات التجارية خدمات في هذا المتسع الفسيح، بما في ذلك WhiteOps[264].

وقد ركّزت الأبحاث وعملية تطوير الأدوات من قبل أطراف ثالثة بشكل أساسي على روبوتات تويتر، بسبب القيود التي تفرضها واجهة برمجة التطبيقات API لفيس بوك. ويُعدّ العنصر التمكيني الرئيسي لهذه المشروعات البيانات حول روبوتات الويب المثبتة والحسابات الدمى (أي هُويّات مزورة على الإنترنت يتم تشغيلها من قبل البشر) وبيانات من وسائل التواصل الاجتماعي الخاصة بها (على سبيل المثال التدوينات، وملف التعريف الاجتماعي، والمشاركات ونقرات الإعجاب). وتماماً كما هي الحال مع جميع عمليات التعلم الآلي، تُعد هذه البيانات ضرورية لتدريب الخوارزميات على الكشف عن روبوتات الويب والحسابات الدمى. هذا وقد أنشأ الكثير من مجموعات البيانات هذه أكاديميون (على سبيل المثال DARPA Twitter Bot Challenge (Subrahmanian et al., 2016) وBot Repository[265]. وحده تويتر قد نشرت علناً حتى اليوم مجموعات بيانات هامة[266] لمساعدة الباحثين المستقلين في هذا المجال.

لم تصل الأساليب المتوفرة الناجمة عن الأبحاث الأكاديمية حتى الآن إلى دقة عالية للغاية، لأنها غالباً ما تعمل فقط مع بيانات الحسابات التي يمكن الوصول إليها علناً (مثل وصف الحساب، وصورة ملف التعريف). وقد يتغير هذا الوضع مع قيام فيس بوك وSocial Science One في بداية عام 2020 بنشر مجموعة بيانات واسعة تضمّ عناوين مواقع URL[267] تمت مشاركتها على فيس بوك، بما في ذلك بيانات حول التفاعل وما إذا كانت هذه المشاركات قد وُسمت بسبب خطاب الكراهية أو نتيجةً للتحقق من الوقائع. وغالباً ما تستخدم شركات التواصل الاجتماعي معلومات إضافية مرتبطة بالحساب، بما في ذلك عناوين بروتوكول الإنترنت وتفاصيل عن تسجيل الدخول، وحسابات البريد الإلكتروني، وذاكرات المتصفّح، ممّا يجعل المهمة أسهل إلى حدّ ما. ويصف تويتر عملية الملكية الخاصة به قائلاً «نحن نعمل مع الآلاف من الإشارات والسلوكيات لتوجيه تحليلنا وتحقيقاتنا. علاوةً على ذلك، فإن أياً من إجراءاتنا الوقائية لتحدي

262 http://www.bot-ou-humano.dcc.ufmg.br/

263 http://www.monitor-de-whatsapp.dcc.ufmg.br/

264 https://www.whiteops.com/

265 https://botometer.iuni.iu.edu/bot-repository/

266 https://about.twitter.com/en_us/advocacy/elections-integrity.html#data

267 https://dataverse.harvard.edu/dataverse/socialscienceone

الحسابات الهادفة إلى التلاعب بالمنصات (ما يصل إلى 8–10 مليون حساب أُسبوعياً) مرئية في العينة الصغيرة المتاحة في واجهة برمجة التطبيقات العامة التي تخصّنا» (Roth, 2019).

يقدم عدد من الكيانات التجارية أدوات ذات صلة لتحليل الشبكات (مثل Graphika [268])، في حين تركّز مبادرات مقبلة تموّلها الحكومة في الولايات المتحدة مثل SEMAFOR على تحديد متعدد الوسائط للمعلومات المُضلّلة من خلال استخدام مؤشرات النزاهة المادية والدلالية والبصرية والرقمية.

أدوات لمساعدة الأطراف الثالثة في التحقُّق من الوقائع

تقوم منظمات التحقُّق من الوقائع والشركات الناشئة بتطوير عدد من الأدوات الآلية للتحقّق من الوقائع، مثل FullFact [269]، وReporters Lab التابع لجامعة ديوك[270]، وFactmata [271]، وChequado [272]، وContentCheck [273]، وهي تهدف إلى مساعدة الأشخاص الذين يتحققون من الوقائع في مهامهم، كالكشف التلقائي عن الادعاءات الوقائعية التي يدلي بها السياسيون وغيرهم من الشخصيات البارزة في النصوص التلفزيونية والأخبار على الإنترنت، مثل Full Fact's Live tool [274] وDuke›s Tech&Check [275]، الذي يستخدم Claimbuster (Funke, 2018).

وتتيح أدوات آلية أخرى تتبُّع ذكر ادعاءات كاذبة معروفة بالفعل، مثل أداة Full Fact's Trend، والتحقُّق التلقائي من الادعاءات الرقمية البسيطة مقارنةً بقواعد بيانات موثوقة، مثل Full Fact Live.
تُستكمل هذه الأدوات بقواعد بيانات وجهود الاستعانة بالجمهور من أجل إنشاء قواعد بيانات لمصادر المعلومات المُضلّلة أو الادعاءات الكاذبة وعمليات التحقُّق من الوقائع القائمة. وهي تشمل جهوداً مثل Storyzy [276] التي تضمّ قاعدة بيانات لمواقع الأخبار وقنوات الفيديو الزائفة (30,000 مصدر للمعلومات المُضلّلة بحلول أوائل العام 2020)، WeVerify [277] الذي يقوم ببناء قاعدة من البيانات المتسلسلة تضمّ الادعاءات الكاذبة والمُحتوى الزائف المعروف، وكذلك مواقع مثل Rbutr [278] الذي يوفّر، بدلاً من التحقّق من الوقائع، روابط ينشئها المجتمع إلى صفحات الدحض. وتقوم أدوات التحقُّق من الوقائع الآلية، مثل Full Fact Live [279] وDuke›s Tech&Check [280]، بالتحقُّق أيضاً من الادعاءات الواردة عبر مقارنتها بعمليات التحقُّق من الوقائع المخزّنة إما في قواعد بيانات داخلية و/أو المجمّعة آلياً بناءً على ادعاءات جديرة بالثقة ومتقاسمة علناً تم التحقُّق منها؛ ترتبط بالمخطط القياسي المفتوح لمراجعة الادعاءات Claim Review.

268 https://www.graphika.com/

269 https://fullfact.org/

270 https://reporterslab.org/

271 https://factmata.com/

272 https://chequeado.com/

273 https://team.inria.fr/cedar/contentcheck/

274 https://fullfact.org/automated

275 https://reporterslab.org/tech-and-check/

276 https://storyzy.com/about

277 https://weverify.eu/

278 http://rbutr.com/

279 https://fullfact.org/automated

280 https://reporterslab.org/tech-and-check/

يجري أيضاً البحث في طرق آلية للتحقّق من الوقائع إستناداً إلى معالجة اللغات الطبيعية (NLP) والتقنيات القائمة على الذكاء الاصطناعي. وقد ركّزت إحدى المقاربات الأساسية على تحديد ادعاءات إحصائية بسيطة (على سبيل المثال، يبلغ عدد سكان المملكة المتحدة 60 مليون شخص) والتحقُّق من صحّتها مقارنةً بقاعدة بيانات منظّمة (Vlachos & Riedel, 2015). وفي حين تتحسّن دقة هذه الأساليب بشكل مستمر بفضل إنشاء مجموعات بيانات كبيرة من الادعاءات النصية المشروحة (Thorne, Vlachos et al., 2018)، فإنها لا تزال تُعتبر غير كافية للاستخدام العملي (Babakar & Moy, 2016). ومع ذلك، ونظراً إلى أن المزيد والمزيد من الادعاءات التي تحقق منها الإنسان يتم مشاركتها علناً في أشكال تستطيع الآلة قراءتها، مثل Claim Review، فهي ستساعد خوارزميات التحقُّق من الوقائع استناداً إلى معالجة اللغة الطبيعية والذكاء الاصطناعي على الوصول إلى مرحلة النضج. وفي الوقت الحاضر، وكما لاحظ تقرير معهد رويترز عن التحقُّق الآلي من الوقائع: «يتفق كلّ من الباحثين والممارسين على أن الوعد الحقيقي لتكنولوجيا التحقُّق الآلي من الوقائع في الوقت الحالي يكمن في أدوات تساعد الجهات التي تتحقّق من الوقائع على تحديد الإدعاءات والتحقيق فيها، وتقديم استنتاجاتها بأكبر فعالية ممكنة» (Graves, 2018).

أدوات شبه آلية لاستكمال التدقيق في المُحتوى

يُعنى التدقيق في المُحتوى بالتأكد مما إذا تم التلاعب بصورة أو مقطع فيديو أو «ميم meme» أو إذا كانت تروج لمعلومات مغلوطة. وقد ركزت بعض من أفضل الأدوات المعروفة على التدقيق عبر الاستعانة بالجمهور (مثل: CheckDesk، وVeri.ly)، وصحافة المواطن (مثل: Citizen Desk)، أو مخازن الوقائع/الشائعات التي تم التحقُّق منها (مثل: Emergent، وFactCheck). وتشمل منصات التدقيق ومنتجاته الأكثر نجاحاً في الوقت الحالي SAM [281]، وCitizen Desk [282]، وCheck [283] وTruly Media [284]. وتتوفر أيضاً بعض أدوات المتصفّح والبرامج المساعِدة الموجّهة إلى الصحفيين، مثل: InVID/WeVerify plugin [285]، بالإضافة إلى Frame by Frame [286] (برامج مساعِدة للتحقق من الفيديو)، وVideo Vault [287] (أرشفة الفيديو والبحث العكسي للصور)، وRevEye [288] (البحث العكسي للصور)، وJeffrey's Image Metadata Viewer [289] (التحقُّق من الصور)، وNewsCheck [290] (قائمة تحقق). وتشمل البرامج المساعِدة التي تتيح رصد مُحتوى الويب ووسائل التواصل الاجتماعي Storyful's Multisearch [291] للبحث في تويتر وVine ويوتيوب وتامبلر Tumblr وانستغرام وSpokeo، مع نتائج معروضة في علامات تبويب منفصلة، من دون تحليل عابر للمواد الإعلامية أو الشبكات الاجتماعية؛ وDistill [292] الذي يرصد صفحات الويب.

281 https://www.samdesk.io/

282 https://www.superdesk.org/

283 https://meedan.com/en/check/

284 https://www.truly.media/

285 https://weverify.eu/verification-plugin/

286 https://chrome.google.com/webstore/detail/frame-by-frame-for-youtub/elkadbdicdciddfkdpmaolomehalghio

287 https://www.bravenewtech.org/

288 https://chrome.google.com/webstore/detail/reveye-reverse-image-sear/keaaclcjhehbbapnphnmpiklalfhelgf

289 http://exif.regex.info/exif.cgi

290 https://firstdraftnews.org/latest/launching-new-chrome-extension-newscheck/

291 https://chrome.google.com/webstore/detail/storyful-multisearch/hkglibabhninbjmaccpajiakojeacnaf

292 https://chrome.google.com/webstore/detail/distill-web-monitor/inlikjemeeknofckkjolnjbpehgadgge

فـي مـا يتعلّـق بالتحليـل الجنائـي للصـور ومقاطـع الفيديـو، تتوفّـر مجموعـة مـن الأدوات، مثـل Forensically[293]، وFotoForensics[294]، وImage Verification Assistant[295] التـي تـم تطويرهـا فـي إطـار مشـروع REVEAL FP7 التابـع للاتحـاد الأوروبـي وInVID/WeVerify للتحقـق مـن الصـور والفيديـو[296] (تـرد مناقشـتها فـي مـا يلـي). وتسـتند الوظائـف المعروضـة حاليـاً إلـى خوارزميـات تسـلّط الضـوء علـى المناطـق التـي تـم التلاعـب بهـا، وعلـى تصنيـف البيانـات الوصفيـة وتحليلهـا، والاسـترجاع شـبه المـزدوج المسـتند إلـى مطابقـة الإطـار الرئيسـي مـن خـلال البحـث العكسـي للصـور (عـادةً مـن خـلال غوغـل). إنّ كل هـذه الأدوات محـدودة، خاصـةً عندمـا يتعلـق الأمـر بمراجعـة مـادة إعلاميـة ذات دقـة منخفضـة و/أو تـم ضغطهـا أو مشـاركتها عبـر منصـة واحـدة أو أكثـر مـن منصـات مشـاركة الفيديـو/التواصـل الاجتماعـي. وبالإضافـة إلـى ذلـك، يتطلـب إسـناد التحليـل الجنائـي عـادةً مسـتوى كبيـر مـن المهـارات التقنيـة.

وقـد مـوّل الاتحـاد الأوروبـي، مـن خـلال الإطـار الخامـس لبرنامـج أُفـق 2020 الخـاص بـه، «إجـراءات ابتكاريـة» لمـدة ثـلاث سـنوات، ونشـاطاً فـي حقـل التنسـيق والدعـم يعالـج علـى وجـه التحديـد المعلومـات المُضلّلـة. وتشـمل هـذه المبـادرات مـا يلـي:

يهـدف مشـروع EUNOMIA[297] إلـى إنشـاء رفيـق لوسـائل التواصـل الاجتماعـي فـي كل مـن نسـخ الهواتـف المحمولـة والحاسـوب لمسـاعدة المسـتخدمين فـي التعـرّف إلـى مسـتخدم وسـائل التواصـل الاجتماعـي الـذي هـو المصـدر الأصلـي لمعلومـة مـا، وكيفيـة انتشـارها، وكيفيـة تعديلهـا فـي سلسـلة مـن المعلومـات، ومـدى احتمـال أن تكـون هـذه المعلومـة جديـرة بالثقـة. وسـيتم اختبـار تكنولوجيـا EUNOMIA فـي حـالات جديـدة تـم إنشـاؤها خصيصـاً مـن قبـل منصـة Mastodon للمدونـات الصغيـرة وشـبكة التواصـل الاجتماعـي Diaspora، مـع مشـاركة مسـتخدمين فـي التقييـم التجريبـي. ويضـم كونسورسـيوم EUNOMIA عشـرة شـركاء مـن 9 بلـدان فـي الاتحـاد الأوروبـي.

مـن جهتـه، يسـعى مشـروع Provenance[298] إلـى تمكيـن المواطنيـن مـن تقييـم المُحتـوى عبـر الإنترنـت، وإلـى تطويـر كفـاءات الإلمـام الرقمـي. وفـي الوقـت نفسـه، يخطـط مشـروع Provenance لمسـاعدة منشـئي المُحتـوى علـى حمايـة عملهـم الأصلـي مـن سـوء الاسـتخدام والتلاعـب، مـن خـلال تسـجيل العمـل الأصلـي فـي سـجل قواعـد البيانـات المتسلسـلة، وتتبـع كيفيـة انتشـاره، وتحديـد أي تلاعبـات تحـدث لاحقـاً. ويجمـع كونسورسيوم Provenance ستة شركاء من أربع دول في الاتحاد الأوروبي.

يركّـز مشـروع Social Truth[299] علـى إنشـاء نظـام بيئـي مفتـوح ومـوزّع وخدمـات تدقيـق فـي المُحتـوى للتحقـق مـن مصـادر المعلومـات أثنـاء عمليـة الإنتـاج، وتوفيـر رفيـق رقمـي (روبـوت الويـب للدردشـة) للمسـاعدة فـي التحقُّـق مـن المُحتـوى، بالإضافـة إلـى تصنيفـات محـركات البحـث ومنـع المواقـع الاحتياليـة مـن الإعـلان.

293 https://29a.ch/photo-forensics/#forensic-magnifier
294 http://fotoforensics.com/
295 http://reveal-mklab.iti.gr/reveal/
296 https://weverify.eu/verification-plugin/
297 https://www.eunomia.social/
298 https://www.provenanceh2020.eu/
299 http://www.socialtruth.eu/

وللكشف عن المعلومات المُضلّلة، يستخدم مشروع Social Truth كل من تقنية الذكاء الاصطناعي والثقة الناجمة عن التدقيق في المُحتوى وسلامته استناداً إلى تقنية قواعد البيانات المتسلسلة. ويجمع كونسورسيوم Social Truth 11 شريكاً من ستة بلدان في الاتحاد الأوروبي.

تهدف أداة WeVerify [300] (المذكورة أعلاه) إلى تطوير أساليب وأدوات ذكية للتحقّق من المُحتوى وتحليل المعلومات المُضلّلة؛ تشمل الإنسان. وسيتم تحليل وسائل التواصل الاجتماعي ومُحتوى الويب وإدراجهما في سياقهما داخل النظام البيئي الأوسع على الإنترنت، من أجل كشف المُحتوى الملفّق من خلال التدقيق في المُحتوى العابر للوسائط، وتحليل الشبكات الاجتماعية، وكشف الزيف الموجّه بدقة، وكشف الصور الزائفة الفجّة deep fakes وقاعدة بيانات عامة للمنتجات الزائفة المعروفة تقوم على قواعد البيانات المتسلسلة. ويتم دمج أدوات WeVerify في أداة Truly Media (وهي أداة تدقيق تجارية) وفي برنامج التدقيق المساعِد InVID/WeVerify، وهي صندوق أدوات مفتوحة المصدر للتدقيق يُستخدم على نطاق واسع من قبل الجهات التي تتحقّق من الوقائع. وتجمع أداة WeVerify سبعة شركاء من ست دول في الاتحاد الأوروبي.

سوما SOMA [301] هو إجراء للتنسيق والدعم، أنشأ مرصداً اجتماعياً للتضليل ولتحليل وسائل التواصل الاجتماعي هدفه دعم الباحثين والصحفيين والجهات التي تتحقّق من الوقائع في نضالهم ضد المعلومات المُضلّلة. تندرج في صميم «مرصد سوما» SOMA للتضليل منصة تعاونية على شبكة الإنترنت (Truly. Media) للتحقق من المُحتوى الرقمي (الذي ينتجه المستخدم) وتحليل انتشاره في النقاش الاجتماعي. يهدف صندوق أدوات ذو صلة اسمه DisInfoNet Toolbox إلى دعم المستخدمين ليفهموا حركيّات نشر الأخبار (الزائفة) عبر وسائل التواصل الاجتماعي وتعقّب منشأ المعلومات الكاذبة وناشريها. يجمع «مرصد سوما» SOMA خمسة شركاء من ثلاثة بلدان.

بدأ مشروع Fandango [302] قبل عام واحد من المشروعات السابقة في 2019 ويستمر حتى نهاية عام 2020. يهدف هذا المشروع إلى آلية عملية الكشف عن المعلومات المُضلّلة والتحقُّق من الوقائع من خلال تحليل البيانات الضخمة والمقاربات اللُّغوية والشبكية. ويخطط المشروع لبناء علامات لمصداقية المصدر ووحدة لملفات التعريف، ووحدة للكشف عن الرسائل المُضلّلة، وكاشف للتزييف، وأدوات للكشف عن النسخ– التحريك لتحليل الصور والفيديو، ووحدة لتحليل الرسم البياني الاجتماعي. ويجمع المشروع ثمانية شركاء من خمسة بلدان.

تُواصل الحكومة الأميركية، عبر برنامج orDARPA MediF [303] (وكذلك عن طريق تحديات التحليل الجنائي الإعلامي من NIST [304]) الاستثمار في مجموعة من مقاربات التحليل الجنائي اليدوي والآلي، التي تشمل التحسينات على المقاربات القائمة على أساس التناقضات في JPEG/MPEG لتحديد متى تم نسخ/لصق عناصر أُخرى داخل صورة أو ما إذا تم قطع عنصر من ملف صورة أُخرى. وتشمل أيضاً تتبع معرّفات

....

300 https://weverify.eu/

301 https://www.disinfobservatory.org/

302 https://fandango-project.eu/

303 https://www.darpa.mil/program/media-forensics

304 https://www.nist.gov/itl/iad/mig/media-forensics-challenge-2018

الكاميرا على أساس PRNU (وهو مقياس لاستجابة كل خلية للضوء في مجموعة أجهزة الاستشعار في الكاميرا، يوفر «بصمة» فريدة من نوعها للكاميرا عند التقاط صورة). تتداخل بعض هذه المقاربات مع مقاربات المصدر الوارد ذكرها في القسم 7.3 – على سبيل المثال، تترك أداة eWitness لتتبع المصدر آثاراً جنائية تم تصميمها كجزء من التكنولوجيا الخاصة بها (Newman, 2019a)، في حين أن بعض الشركات الناشئة التي تُعنى «بالتقاط الصور الخاضع للسيطرة controlled capture» تستخدم تقنية الرؤية الحاسوبية computer vision (التقنيات العلمية المتعلقة بتحديد الصورة وتصنيفها) للتحقّق من وجود أدلة على إعادة التقاط صورة موجودة.

إن معظم الخوارزميات التي هي قيد التطوير في برامج مثل برنامج DARPA Medifor وغيرها من برامج تمويل التحليل الجنائي الإعلامي ذات الصلة لم تتوفّر حتى الآن كأدوات يتفاعل معها المستخدم user-facing. من جهته، وضع Jigsaw التابع لشركة Alphabet أداة Assembler، وهي أداة ألفا alpha tool[305]، في متناول صحفيين مختارين في أوائل عام 2020. تتيح هذه الأداة إجراء تحليل جنائي إعلامي تقليدي، والكشف عن الوجوه الاصطناعية أو المركّبة التي استحدثتها أداة تُعرف باسم StyleGan.

تميل بعض الأدوات الأكثر دقةً إلى الجمع بين البيانات الوصفية والتفاعلات الاجتماعية والإشارات البصرية والملف التعريفي للمصدر (أي وكيل المنشأ)، وغيرها من المعلومات السياقية المحيطة بصورة أو شريط فيديو لمساعدة المستخدمين في مهمة التدقيق في المُحتوى. وتتماشى هذه المقاربات الدلالية بشكل وثيق مع الطريقة التي يستخدم فيها الصحفيون والمحققون أدوات OSINT (أي المعلومات الاستخباراتية مفتوحة المصدر) وممارسات التحقُّق البصري. ومن بين الأدوات هذه الأكثر استخداماً، يمكن ذكر البرنامج المساعِد Invid/WeVerify[306] (Teyssou et al., 2017)، وأداة التحقُّق من الفيديو على الإنترنت Youtube Data Viewer[307] التابعة لمنظمة العفو الدولية. وتتيح هذه الأداة الأخيرة استخراج قوائم البيانات الوصفية وإجراء بحث عن التشابه على أساس الصور باستخدام إطارات رئيسية keyframes.

أدوات للكشف عن أشكال جديدة من المواد الإعلامية التي تم التلاعب بها عبر الخوارزميات

لا تتوفر حتى اليوم أدوات متاحة تجارياً للكشف عن مجموعة واسعة من الأشكال الجديدة للمواد الإعلامية السمعية– البصرية التي تم التلاعب بها بواسطة الذكاء الاصطناعي والمعروفة باسم «الصُّور الزائفة الفجّة» deepfakes و/أو «المواد الإعلامية المركّبة». على ذلك، لم تكشف المنصات عن طبيعة الأدوات التي تنشرها.

على الرغم من ذلك، يتم تطوير الكثير من المقاربات، بما في ذلك عدد منها يعتمد إما على تطوير إضافي للتحليل الجنائي الإعلامي، أو يستخدم أشكال الشبكات العصبية نفسها التي كثيراً ما تُستخدم لإنشاء صور زائفة فجّة ولكن في إطار عملية الكشف. وتستخدم أشكال أخرى من الكشف التعلم الآلي، ولكنها تعتمد على تقنيات التشكيك في التكامل الدلالي للصور والقصص لتحديد التلاعب (Verdoliva, 2020).

305 https://jigsaw.google.com/assembler/

306 https://weverify.eu/verification-plugin/

307 https://citizenevidence.amnestyusa.org/

تستطيع مقاربات الكشف باستخدام تقنيات الاستحداث القائمة على شبكات الخصومة التوليدية الجديدة (GAN)– التي تُستخدم لتصميم صور زائفة فجّة ومواد إعلامية مُركّبة أُخرى– أن تستخدم نفس المقاربة التقنية لتحديد الصور الزائفة (Gregory, 2019). وفي بداية عام 2020، صدرت الأدوات الأولى في إطار Jigsaw Assembler المُشار إليه أعلاه. وينبغي، في هذا الصدد، أن نتوقّع أن يدخل البعض الآخر قريباً إلى السوق ويكون متاحاً للصحفيين، إما كبرامج مساعِدة plug-ins أو كأدوات على المنصات في عام 2020. وستعتمد هذه الأدوات بشكل عام على الحصول على بيانات تدريبية (أمثلة) عن مقاربة التزوير، لذا لن تكون بالضرورة فعالة على أحدث أساليب التزوير. على سبيل المثال، تستحدث مشروعات التحليل الجنائي، مثل ++FaceForensics، صوراً زائفة باستخدام أدوات مثل FakeApp ومن ثم تستعمل هذه الكميات الكبيرة من الصور الزائفة كبيانات تدريبية للشبكات العصبية التي تجري تمارين الكشف الوهمية (Rossler et al., 2018). غير أن الشركات الكُبرى بدأت هي أيضاً بالاستثمار في دعم الأبحاث المستقلة، فضلاً عن زيادة مجموعات بيانات لتسهيل تطوير الحلول. ومن الأمثلة في هذا السياق، عمل غوغل مع مشروع Face Forensics (Dufour & Gully, 2019)، والملف الصوتي المركّب synthesised audio (Stanton, 2019)، بالإضافة إلى تحدّي الكشف عن الصور الزائفة الفجّة Deepfakes Detection Challenge (Schroepfer, 2019) الذي أطلقه فيس بوك ومايكروسوفت وأمازون والشراكة حول الذكاء الاصطناعي ومجموعة من الأكاديميين.

تنظر المقاربات الأخرى في هذا المجال أيضاً في التطورات التي شهدها التحليل الجنائي الإعلامي لتحديد توقيعات الصور المميزة للمواد الإعلامية التي تم استحداثها من خلال شبكات الخصومة التوليدية GAN (Marra et al., 2018) (على غِرار «بصمات» PRNU الخاصة بالكاميرات التقليدية). وخارج إطار برامج مثل شراكة DARPA MediFore، يعمل عدد من الشركات التجارية والمؤسسات الأكاديمية في مجال الكشف القائم على شبكات الخصومة التوليدية، بما في ذلك (وليس حصراً) DeepTrace Labs[308]، وFaculty AI[309] وWeVerify[310] ومعهد روتشستر للتكنولوجيا[311]. وتتمحور الأسئلة الرئيسية بشأن هذه الأدوات حول مدى نجاحها مع أنواع مختلفة من التلاعب، ومدى صلابتها مع تطوّر عمليات التزوير وتحسّنها، وكيف ستقدم نتائجها بطرق قابلة للتفسير ومفيدة للصحفيين والمستخدمين. ويقدم التقرير الأخير الصادر عن اللجنة التوجيهية حول النزاهة الإعلامية التابعة للشراكة حول الذكاء الاصطناعي، والذي أشرف على تحدّي الكشف عن الصور الزائفة الفجّة Deepfakes Detection Challenge، المزيد من الإرشادات حول كيفية تفعيل هذه المخاوف لتطوير تقنيات الكشف (Partnership on AI, 2020).

تشمل الأشكال الجديدة للتحليل الجنائي اليدوي والمُميكن مقاربات تعتمد على الفهم الحالي لكيفية الكشف عن التلاعب في الصور والنسخ/اللصق/التحرير، فضلاً عن مقاربات متطورة مكيّفة للصور الزائفة الفجّة، مثل استخدام التحليل الطيفي لتحديد خصائص مميزة للكلام الاصطناعي[312]، أو فكرة استخدام مؤشرات بيولوجية[313] للبحث عن تناقضات في الصور الزائفة الفجّة (AlBadawy et al., 2019). وقد تم أيضاً اقتراح مجموعة من المقاربات لإنشاء ما يُسمّى بـ«الخصائص البيومترية الناعمة»

....

308 https://deeptracelabs.com/

309 https://faculty.ai/

310 https://weverify.eu/

311 https://aiethicsinitiative.org/news/2019/3/12/announcing-the-winners-of-the-ai-and-the-news-open-challenge

312 https://www.researchgate.net/publication/333393640_Detecting_AI-Synthesized_Speech_Using_Bispectral Analysis

313 https://www.researchgate.net/publication/333393640_Detecting_AI-Synthesized_Speech_Using_Bispectral Analysis

للشخصيات العامة الرئيسية، مثل المرشحين للرئاسة الأمريكية لعام 2020، وهي تتيح التحقُّق من صور زائفة فجّة يشتبه بها لمعرفة ما إذا تمت محاكاة الصوت وحركة الشفتَين (Agarwal & Farid, 2019; Beavers,2019). ففي المُحتوى الأصيل، يجب أن يقوم ارتباط بين ما يقوله الشخص وطريقة قوله (نمط مميّز لحركات الرأس يرتبط بكيفية لفظ أفراد معروفين كلمات معينة).

تنظر مقاربات أخرى في مسائل التكامل المادي (‹هل يتناقض الأمر وقوانين الفيزياء؟›)، مثل ضمان عدم وجود أي تناقض في الإضاءة والانعكاس والصوت، إضافةً إلى مراجعة التكامل الدلالي للمشاهد («هل الأمر منطقي؟»)، والنظر في مقاربات التحليل الجنائي السمعي[314] لتحديد عمليات التزوير، وتحديد مصدر الصورة وأصلها (Moreira, et al., 2018).

وتتعلق مقاربات مُميكنة أخرى لتتبُّع الصور الزائفة الفجّة بالنظم المميكنة الحالية للكشف عن المُحتوى المتوفر على المنصات، بما في ذلك المقاربات القائمة على أصل الصور ومصدرها. ترتبط المقاربات التي تركّز على مصدر الصور بشكل وثيق بمحركات البحث عن الصور القائمة التي تستخدم البحث العكسي للصور أو عمليات بحث لأوجه الشبه الأخرى لتحديد نسخ سابقة أو مشابهة لصورة ما. أما مقاربات نسل أو أصل الصور، فتعتمد على فهارس مماثلة للصور الموجودة للبحث عن تاريخ عناصر الصورة واكتشاف إعادة استخدام العناصر داخل الإطار.

تتضمّن أدوات الكشف المميكن عن النص المُستحدث بواسطة الذكاء الاصطناعي أداة Grover[315] (Zeller et al., 2019) أو نموذج Glitr[316] (Strobelt & Gehrmann, 2019). وأداة Grover هي نظام تزويدي فضلاً عن كونه نظام كشف. وتماماً كما هي الحال مع مقاربات أخرى قائمة على التعلّم المتقدم، تُعدّ هذه الأدوات عموماً أقل صلابةً عند تطبيقها على نص تم إنتاجه بواسطة نماذج وبيانات مختلفة عن تلك التي تدرّبت عليها الأدوات. وقد واصل المطورون الأوائل للأساليب ومجموعات البيانات في هذا المجال –على سبيل المثال -Open AI's GPT-2 (Solaiman et al., 2019) نشر معلومات عن الرموز والوزن النموذجي لتسهيل الكشف عن مخرجات النماذج المستمدة من GPT-2. وتقوم جهات فاعلة تجارية تعمل على مُكافحة التّضليل وعلى التحقيق (كما أُشير إليه في 4.2) بالتحقيق في فائدة الكشف عن نص تم إنتاجه آلياً (Rahman et al., 2019).

4.2.6 من هي الجهات الفاعلة الرئيسية ومن يموّل هذه الاستجابات؟

تدعم شركات الإنترنت القائمة، من خلال نماذجها التجارية (مثل الإعلانات الهادفة)، الاستجابات الداخلية كما توفّر بعض الخدمات لأطراف ثالثة. وتشمل هذه الخدمات موارد مسجّلة الملكية، مثل الكشف المميكن عن روبوتات الويب، والموارد المقيدة كالمعلومات المؤمنة لأطراف ثالثة تتحقّق من الوقائع، ومجموعات البيانات للكشف عن الصور الزائفة الفجّة. وفي بعض الحالات، تتوفّر قُدرات مُتاحة للجمهور، كالبحث عن أوجه الشبه أو البحث عن الصور. وبشكل عام، تتوفّر هذه الخدمات بشكل مجاني.

[314] https://newsinitiative.withgoogle.com/dnifund/dni-projects/digger-deepfake-detection/

[315] https://grover.allenai.org/

[316] http://gltr.io/

هناك مقاربات أخرى، لا سيّما في ما يتعلق بأدوات الأطراف الثالثة، تكون مزيجاً من المقاربات الممولة من الحكومات/المنح المشروطة (مثل صناديق DARPA والاتحاد الأوروبي لمقاربات الكشف والتحقُّق)، فضلاً عن مبادرات لا تهدف للربح وشركات ناشئة.

5.2.6 دراسة حالة عن الاستجابة: معلومات مضللة حول جائحة فيروس كورونا

تتمثّل إحدى التبعات التقنية والخوارزمية الرئيسية لجائحة فيروس كورونا في انتقال شركات الإنترنت الكبرى إلى مراقبة مميكنة إضافية للمُحتوى وتسامح موصوف أكبر مع الإيجابيات الكاذبة. وعلى الرغم من أن هذه المسألة توجّهها قضايا متعلقة بالصحة في مكان العمل وأمن المعلومات، مع انتقال اليد العاملة (من موظفين ومتعاقدين) إلى العمل عن بُعد، غير أنها تُعتبر تجربة تتناول عملية مميكنة بشكل أكبر لمراجعة المُحتوى. ويشير فيس بوك إلى أنه مع «تقليص الأيدي العاملة وتأديتها عملها عن بُعد، سوف نعتمد اليوم اعتماداً أكبر على أنظمتنا المميكنة للكشف عن المُحتوى المخالف وإزالته وتعطيل الحسابات. ونتيجةً لذلك، نتوقع ارتكاب المزيد من الأخطاء، وسوف تستغرق المراجعات وقتاً أطول من المعتاد». ويُشير فيس بوك أيضاً إلى أنه «عادةً عندما نزيل المُحتوى، نقدّم للشخص الذي نشره خيار طلب مراجعة المُحتوى مرةً أخرى إن كان يعتقد أننا أخطأنا. أما الآن، ونظراً إلى انخفاض الأيدي العاملة لدينا، سنمنح الأشخاص خيار إخبارنا بأنهم لا يوافقون على قرارنا وسنراقب ردود الأفعال هذه لتحسين دقتنا في أداء المهام، ولكننا لن نراجع على الأرجح المُحتوى مرةً ثانيةً[317]». وقد تحلّت شركات أخرى أيضاً بالمصداقية إزاء عواقب التحول إلى المزيد من الميكنة. فيشير غوغل مثلاً إلى أن «أنظمتنا المميكنة قد لا تصنّف دائماً المُحتوى بدقة لأغراض إزالته، وقد تكون المراجعة البشرية لهذه القرارات أبطأ[318]». من جهته، يذكر تويتر أن الأمر: «يزيد من استخدامنا للتعلم الآلي والميكنة للاضطلاع بمجموعة واسعة من الإجراءات بشأن المُحتوى الذي يحتمل أن يكون مسيئاً واستغلالياً. نريد أن نكون واضحين: بينما نعمل على ضمان اتساق أنظمتنا، فإنها قد تفتقر في بعض الأحيان إلى السياق الذي تأتي به فرقنا. وقد يؤدي هذا الأمر إلى ارتكاب الأخطاء». ويشير يوتيوب[319] إلى أنّ «الأنظمة المميكنة ستبدأ بإزالة بعض المُحتوى دون مراجعة بشرية، حتى نستمرّ في التصرف بسرعة لإزالة المُحتوى المخالف وحماية نظامنا البيئي، في حين تتوفّر لدينا الحماية في مكان العمل... وأثناء قيامنا بذلك، قد يرى المستخدمون والمبدعون ارتفاعاً في عدد مقاطع الفيديو المحذوفة، بما في ذلك بعض مقاطع الفيديو التي قد لا تنتهك السياسات[320]».

وقد أشارت إحدى الدراسات التي أُجريت في منتصف عام 2020 إلى الصعوبة التي يواجهها فيس بوك في التعامل مع مستويات غزيرة من المعلومات الصحية المُضلّلة على الموقع، بحجة أن الشركة تحتاج إلى تحسين استجاباتها الخوارزمية (Avaaz 2020). وقد وجدت الدراسة، على وجه الخصوص، أن 16٪ فقط من المُحتوى الذي حدده الباحثون على أنه معلومات مضللة ذو صلة بالصحة يحمل علامة تحذير. وقد شوهد مُحتوى صحي مغلوط ومضلل 3.8 مليار مرة في الأشهر الإثني عشر الماضية، وقد بلغ الأمر ذروته خلال جائحة فيروس كورونا، وفقاً للبحث (Avaaz 2020).

....

317 https://about.fb.com/news/2020/03/coronavirus/

318 https://blog.google/inside-google/company-announcements/update-extended-workforce-covid-19

319 https://blog.twitter.com/en_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html

320 https://youtube-creators.googleblog.com/2020/03/protecting-our-extended-workforce-and.html

تنطوي المراقبة المميكنة للمُحتوى على خطرَين واضحَين، وهما أنه في غياب المراجعة البشرية ذات الصلة، تستحدث المراقبة هذه إيجابيات كاذبة مستمرة لانتهاكات سياسة المُحتوى، وأن الحق في الطعن في القرارات لأمر ضروري. يعلّق أحد المراقبين قائلاً: «مع توقّف الكثير من مراقبي المُحتوى البشر فجأةً عن العمل، اضطرت المنصات إلى الاعتراف بالحدود الفعلية لتكنولوجيتها... فمن المستحيل مراقبة المُحتوى على نطاق واسع بشكل مثالي – ويجب على المنصات اتخاذ الملايين من القرارات يومياً ولا يمكنها اتخاذ القرار الصحيح في كل مرة. وبما أن الخطأ أمر لا مفر منه، يتطلّب تصميم نظام مراقبة المُحتوى اختيار أنواع الأخطاء التي سيقترفها النظام. وفي سياق الجائحة، عندما أعلنت منظمة الصحة العالمية عن «مشكلة تضليل» ولم يكن بإمكان مراقبي المُحتوى البشر ببساطة الذهاب إلى العمل، اختارت المنصات أن تخطئ من ناحية الإيجابيات الخاطئة وأن تزيل المزيد من المُحتوى» (Douek, 2020).

تعترف الشركات بأن النظم المميكنة الحالية ليست بديلاً عن الرقابة البشرية، وهذا يعزّز الحاجة إلى نظم تصحيح وطعن قوية، كما أشار إليه مقرر الأمم المتحدة الخاص المعني بتعزيز وحماية الحق في حُريّة الرأي والتّعبير (Kaye, 2018)). ويشير المراقب نفسه الذي ذُكر أعلاه إلى ما يلي: «إنّ مراقبة المُحتوى أثناء هذا الوباء هو نسخة مبالغ فيها عن مراقبة المُحتوى طوال الوقت: فالمنصات تقيم توازناً بين مصالح متنوعة عندما تضع قواعدها، وتتخذ خيارات لها تبعات بشأن التفضيلات الخاصة بالأخطاء عند تطبيقها. وينبغي الترحيب بشفافية المنصات غير المعهودة (وإن كانت لا تزال محدودة للغاية) حول هذه الخيارات في سياق الوباء –غير أنه يتعيّن توسيع نطاقها في المستقبل. وينبغي عدم اتخاذ هذه الأنواع من الخَيارات في الخفاء» (Douek, 2020).

6.2.6 كيف يتم تقييم الاستجابات التقنية والخوارزمية؟

مما لا شكّ فيه أن الافتقار إلى البيانات يعيق التدقيق الخارجي في مدخلات معظم العمليات الخوارزمية الداخلية ونماذجها ومخرجاتها داخل المنصات. ويؤدي ذلك أيضاً إلى الحد من قدرة الجمهور على تقييم الخوارزميات الخارجية وخوارزميات الأطراف الثالثة، لأنّ الغرباء لا يستطيعون النفاذ إلى جميع البيانات داخل منصة محددة، ولا إلى بيانات ذات صلة بالسياق تتناول ظاهرة يجب دراستها أو تحديدها. كما أنّ أفراد الجمهور لا يستطيعون الوصول إلى البيانات العابرة للمنصات لتتبع المعلومات المُضلّلة بشكل مناسب. ويعيق هذان العاملان كلاهما التقييم الفعّال.

كما ذُكر أعلاه، يحول غياب الشفافية الأعمق بشأن استخدام النظم الخوارزمية، أو بشأن تنفيذ العناية الواجبة في مجال حقوق الإنسان، دون إجراء تقييم خارجي لفعاليتها في التصدي للتضليل أو تأثيرها على حُريّة التّعبير وحقوق أخرى (Llansó et al., 2020; Gorwa et al., 2020). وتقدم تقارير الشفافية أرقاماً مجمّعة عن الإنفاذ المتعلّق بالحسابات الزائفة[321] مثلاً، ولكنها لا تورد أي تفاصيل.

يتم تقييم نماذج الكشف عن الصور الزائفة الفجّة –استناداً إلى التحليل الجنائي والتعلم المتقدم على حد سواء– على أساس المعايير القياسية ومجموعة اختبار مكوّنة من صور مماثلة ليست جزءً من بيانات التدريب، ولكنها حالياً غير مُختبرة في سياق الاستخدام واسع النطاق للصور الزائفة الفجّة أو الصُّور الإعلامية الاصطناعية التي تم استحداثها باستخدام طائفة واسعة من المقاربات القائمة والجديدة.

[321] https://transparency.facebook.com/community-standards-enforcement#fake-accounts

7.2.6 التحديات والفرص

تعتبر شركات الإنترنت أن مقاربات التعلّم الآلي الرامية إلى التعرّف إلى المعلومات المُضلّلة ومراقبتها تستفيد من إمكانية تنفيذها على نطاق واسع وبسرعة أقرب إلى الوقت الحقيقي من الرقابة البشرية. ويمكن أن توفّر هذه المقاربات آلية لفرز المُحتوى وتوفير آراء متبصّرة للعنصر البشري المسؤول عن الاستجابة للتهديدات داخل الفرق الهامة في الشركات، فضلاً عن الفرق الكبرى (الداخلية والمتعاقد معها من الباطن) المسؤولة عن مراقبة المُحتوى (والمقدّرة بنحو 15,000 اعتباراً من آذار/مارس 2019 في فيس بوك) (Newton, 2019b). وقد لا يتواءم هذان الهدفان بالضرورة مع النتائج المرغوب فيها اجتماعياً في مجال حُريّة التّعبير.

فبصفتها استجابات خوارزمية، تخضع لكلّ من التحيز الضمني والصريح المحتمل في تصميمها وفي بيانات التدريب المُستخدمة لتطويرها (انظر المزيد من المعلومات أدناه). ولكن على مستوى عنصر مُحتوى معيّن، تتعرّض بدرجة أقل للضغوط التي تمارسها الدول وغيرها على المشغّلين البشر الأفراد داخل شركة ما لاتخاذ إجراءات بشأن معلومات مضللة مزعومة.

أما بالنسبة إلى الأطراف الثالثة، بما في ذلك الجهات التي تتحقّق من الوقائع والصحفيين والمحققين الآخرين، فتوفّر الأدوات التي تستخدم التعلم الآلي آليات إضافية لفهم المُحتوى وتسريع عملية صنع القرار، ولكنها تخضع لقيود ناجمة عن غياب سياق إضافي متاح لشركات المنصات. ويمكن أيضاً استخدام هذه الأدوات لتحليل التدابير المُميكنة التي تم تطبيقها بشكل خاطئ أو أُسيء تطبيقها ولتقييم تأثيرها على حُريّة التّعبير.

غير أنّ الأدوات الحالية ليست مناسبة لتحديد المعلومات المُضلّلة على نطاق واسع وفي الوقت الحقيقي وبدقة عالية للغاية. فالاستجابات الخوارزمية داخل المنصات تعاني من مجموعة من الخصائص التي تمسّ حُريّة التّعبير. وبعض هذه الخصائص إجرائية، فيما يعود بعضها الآخر إلى حدود المعايير التقنية لأنظمة الذكاء الاصطناعي، ويتمثّل البعض الآخر في القرارات المتخذة لأسباب تتعلّق بالملكية أو لحماية الأنظمة من اختراق خصمي (Duarte & Llanso, 2017; Llanso et al., 2020). إلى جانب ذلك، يخدم بعضها سياسات لا تراعي المعايير الدولية لحُريّة التّعبير من حيث طريقة إصدار الأحكام بشأن الضرر المحتمل والتناسب. وتبرز تعقيدات إضافية متى كانت السياسات أكثر غموضاً أو أوسع من القانون الدولي لحقوق الإنسان في تعريفها للإرهاب وخطاب الكراهية والتحريض على الأذى (Article 19, 2018a; Article 19, 2018b).

وتُعاني المقاربات القائمة على الذكاء الاصطناعي ممّا يُسمى مشكلة التحيّز التي تحدث في مراحل متعددة من تصميم النظام المميكن وبنائه وتنفيذه (Hao, 2019). وهي تشمل تحديات متعلقة بكيفية تأطير مشكلة ما (مثل: تعريف المعلومات المُضلّلة أو إدراج معايير حقوق الإنسان)، على مستوى جمع البيانات عندما يتم مثلاً جمع بيانات تدريب غير تمثيلية أو ضعيفة التصنيف أو غير ملائمة (أو تحتوي على تحيُّز ضمني أو صريح لصالح مجموعة معينة، كما هو الحال مع الذكاء الاصطناعي في سياقات أخرى)، وعلى مستوى إعداد البيانات لضمان تركيز الخوارزمية على الخصائص البارزة لهدف النظام المُميكن.

تعمل معظم الأدوات بشكل أفضل عندما يتم تدريبها وتطبيقها في مجالات محددة، علماً أنه لا يمكن بالضرورة تطبيقها بنفس الموثوقية عبر سياقات متباينة. ولا تنطبق النظم القائمة على الذكاء الاصطناعي بشكل جيد بين سياقات متنوعة، لا سيّما عندما تكون البيانات التدريبية غير مناسبة لتدريب نماذج التعلم الآلي. ويمكن أن يؤدي ذلك إلى إضعاف فعالية أنواع معينة من المُحتوى –على سبيل المثال المُحتوى الذي تنتجه الأقليات أو اللُّغات التي تتوفّر بشأنها بيانات غير كافية أو بيانات ضعيفة المصدر، كما حدث عند تقييم فعالية التعرّف إلى خطاب الكراهية باللُّغة البورميّة (Stecklow, 2018)، أو الاستهداف المفرط لأنواع معينة من المُحتوى. ويواجه السكّان المهمشون أصلاً المزيد من التهميش بسبب النظم المميكنة. ومما لا شكّ فيه أن هذه القضايا، لا سيما لفائدة فهم المجتمعات واللُّغات الأقل بروزاً، تترك تبعات على أنظمة الذكاء الاصطناعي التي تُحلّل الخطاب متى طُبّقت للكشف عن المعلومات المُضلّلة.

بالإضافة إلى ذلك، تُصمّم جميع أنظمة الذكاء الاصطناعي والخوارزميات وتُنفّذ مع مراعاة أهداف السياسة العامة التي قد تتماشى إلى حد كبير أو بدرجة أقل مع اعتبارات حُريّة التّعبير أو قد تنطوي على تحيُّز ضمني أو صريح على مستوى تصميم السياسات وتأطيرها.

على سبيل المثال، يسلّط فيس بوك الضوء على خمس قيم يستخدمها في معايير المجتمع الخاصة به (Bickert, 2019). وتشمل هذه القيم الصوت والأصالة والسلامة والخصوصية والكرامة. وعلى الرغم من أن فيس بوك يشير إلى استخدام «المعايير الدولية لحقوق الإنسان» لإصدار أحكام بشأن القضايا، فإنه لا يقدم تفاصيل دقيقة عن كيفية القيام بذلك. وبدلاً من التحيز الضمني في تصميم خوارزمية ما، تتخذ شركات الإنترنت قرارات صريحة في سياساتها حول كيفية فهمها لحُريّة التّعبير، مع تداعيات متتالية على مستوى تصميم الخوارزميات وغيرها من الأنظمة المميكنة وتطبيقها، وعلى القرار المُتّخذ بشأن ما يحتاج إلى مراجعة بشرية.

يشكّل تعريف المصطلحات تحدياً آخر في مجال تدريب نظم التعلُّم الآلي. ونظراً إلى التحديات في تعريف المعلومات المُضلّلة (والمعلومات المغلوطة) والخلاف بين البشر بشأن التعاريف، فإنّ هذا النقص في الدقة يحول دون بناء مجموعات بيانات قوية. أما في مجال خطاب الكراهية، فعندما يُطلب من الناس شرح المصطلحات العنصرية، يتّضح أنهم يتفقون مع بعضهم البعض في 69٪ من الحالات فقط (Bartlett et al., 2014). وقد اتّضح أن مهمّة التمييز بين التغريدات المهذبة وتلك غير المهذبة أسهل بالنسبة إلى البشر، وقد تراوحت نسبة الاتفاق بين 80٪ و95٪ استناداً إلى لغة التغريدة (Theocharis et al., 2016). وبالمثل، فإنّ الأداء اللاحق دون انتظار 0-day للكاشفات عن الصور الزائفة الفجّة في مواجهة تقنية تزوير جديدة سيكون مُقوّضاً، طالما أن نماذج الكشف لا تُعمّم جيداً لتشمل تقنيات التزوير الجديدة. وتواجه كاشفات الصور الزائفة العميقة أيضاً نقاط ضعف كبيرة من حيث التعامل مع الضّغط وتحويل الترميز المشتركَين بين الشبكات الاجتماعية، والتعامل مع الاضطرابات الاختراقية التي تعطّل الرؤية الحاسوبية. ويبرز أيضاً نقاش كبير حول الطريقة الفُضلى لعرض البيانات المستمدة من أجهزة الكشف عن التلاعب الجنائي لكي يتمكّن الإنسان من قراءتها وتفسيرها (Verdoliva, 2020).

بالنسبة إلى المعلومات المُضلّلة، يمكن التمييز بين الواقع والتلفيق، ولكن تحديد ما إذا كان الأول يُشكّل الحقيقة والثاني هو دائما باطل (ويتميز عن التهكّم أو الخيال، أو أنه في وضع غير معروف حتى الآن)

لَمسألة أكثر تعقيداً بكثير. وتطرح الميكنة في هذا المجال/في بُعد المُحتوى المعين هذا تحديات أكيدة، تماماً كما تفعل بالنسبة إلى ارتباط المُحتوى بهُويّة زائفة وسلوك مُفتعل (سواءً أكان ذلك مُعد له أم لا). ويعقّد المُحتوى السمعي- البصري أيضاً التمييز الأول حتى، حيث تجري إعادة تدوير الكثير من المُحتوى المستخدم في التّضليل أو يُدرج المُحتوى الأصلي في غير سياقه بحيث يمكن أن يكون المُحتوى الأساسي وقائعياً أو صادقاً ولكن التأطير يكون مُلفّقاً.

بالإضافة إلى ذلك، يعمل الكثير من أنظمة التعلم الآلي التي تعتمد على الشبكات العصبية –مثل الكثير من أدوات الكشف عن الصور الزائفة الفجّة والمواد الإعلامية الاصطناعية، وتلك المستخدمة للكشف عن التلاعب بالمواد الإعلامية القائمة بفعالية أكبر –في حركيّة الاختراق الخصمي المستمر مع جهات فاعلة تحاول خداعها (Verdoliva, 2020).

على الرغم من أن المنصات لا توفّر معلومات مفصّلة عن فعالية أدوات الكشف المميكن الخاصة بها، يمكننا أن نستقي الدروس المستفادة من أحدث الأساليب عن مستويات الدقة في معالجة اللُّغات الطبيعية وغيرها من المجالات. فعلى سبيل المثال، تتمتّع الأساليب الحديثة للكشف عن خطاب الكراهية في الأبحاث الأكاديمية حالياً بدقة تتراوح بين 65–70٪ مقارنةً بالكشف البشري باستخدام التعريف ومجموعة البيانات نفسها (Wulczyn, Thain, & Dixon, 2017). غير أنه من الصعب إعطاء رقم ثابت لأن مجموعات البيانات والمهام تختلف اختلافاً شديداً –وتصل أعلى معدلات الدقة التي لوحظت في الدراسات الحديثة إلى 92٪ (MacAvaney et al., 2019). وعلى الرغم من أن شركات الإنترنت تستطيع النفاذ إلى معلومات إضافية غير عامة عن تدوينة معينة (مثل: عنوان المنشأ لبروتوكول الإنترنت)، فإن الخوارزميات لا تزال غير دقيقة بما فيه الكفاية لاستخدامها بطريقة مميكنة بالكامل. وكمثال على ذلك، تم تفعيل خوارزميات الكشف عن خطاب الكراهية في فيس بوك مؤخراً بسبب قسم من إعلان الاستقلال الأمريكي، الأمر الذي أدى إلى حجب التدوينة المعنية آلياً قبل نشرها الأولي (MacGuill, 2018). وستتعرّض معدلات فشل تبلغ 10٪ حتى للتضخيم بسرعة نظراً إلى نطاق عناصر المُحتوى في شبكة اجتماعية معينة. كما أن النظم المميكنة تجتمع في الكثير من الأحيان مع خوارزميات متعددة، ويتأتّى عن ذلك أن الأخطاء قد تتضخّم بسرعة. وإذا ما سُجّلت تحديات جدية في تحديد مراجع خطاب الكراهية من خلال التعلم الآلي، فإن مسألة التقييم المميكن للمعلومات المُضلّلة (بشأن موضوعات مثل: تغيُّر المناخ حتّى) ستكون أكثر تعقيداً.

لهذه القيود على الدقة تبعات، تماماً كما تنتج تبعات عندما (1) تُصنّف المعلومات المُضلّلة عن طريق الخطأ على أنها صحيحة أو يتم تحديد حسابات روبوتات الويب عن طريق الخطأ على أنها بشرية؛ و(2) تُسجّل إيجابيات كاذبة. وبشكل مباشر، تبرز المشكلات عندما تُصنّف المعلومات الصحيحة بشكل خاطئ على أنها معلومات مضللة أو يتم تحديد مستخدمين حقيقيين بشكل خاطئ على أنهم من روبوتات الويب. والاستنتاج هو أن الأدوات المميكنة الحالية ليست مناسبة للتشغيل المستقل دون إشراف بشري أو دون سبل الإنصاف.

والأمر يصحّ بشكل خاص لأن الأنظمة المميكنة الحالية على المنصات تنطوي على نقاط ضعف إجرائية تشمل غياب الرقابة والشفافية بشأن الخوارزميات، بما في ذلك عدم قدرة الغرباء المستقلين على التدقيق في الحالات التي يُسجّل فيها تحيز في التصميم أو بيانات التدريب أو التنفيذ، أو على تقييم

فعالية المقاربة (Ranking Digital Rights 2020). وقد ذُكرت هذه المشكلة أعلاه عند التطرّق إلى تقييم المقاربات التي تعتمدها الشركات لتحليل الرسائل والجهات الفاعلة والسلوك.

ويشير انعدام الشفافية هذا أيضاً إلى أن الحذف الخاطئ للمُحتوى أو خفض ترتيب المُحتوى أو الجهات الفاعلة يجتمع مع عدم تفسير القرارات الفردية والجماعية لتصنيف المُحتوى ضمن فئة ما، كالمعلومات المُضلّلة. فحتى المحاولات الآيلة إلى التصدي لموضوع مراقبة المُحتوى من خلال إشراف أكثر استقلالية (مثل مجلس الرقابة التابع لفيس بوك Oversight Board, 2019e) لا تُتيح تغيير الخوارزميات الأساسية. وبشكل مماثل، يحول غياب الشفافية الخوارزمية أو الشرح دون استفادة الصحفيين/الجهات التي تتحقّق من الوقائع من تفسير القرارات بشأن المُحتوى.

على ضوء هذه المسائل، تتوفّر مجموعة من المبادئ التي تُتيح تعزيز الشفافية –بما في ذلك مبادئ سانتا كلارا[322] التي تركّز على الأرقام والإشعارات والطعن. ويمكن أيضاً في هذا الصدد ذكر توصيات المُقرِّر الخاص للأمم المتحدة «دايفيد كاي» (Kaye, 2018) بشأن مساءلة شركات الإنترنت، فضلاً عن «مبادرات الشفافية التي تشرح وقع الميكنة والمراقبة البشرية للمُحتوى، ووسم المستخدم أو الوسم الموثوق به لشروط الخدمة».

ويتمثل أحد التحديات الأساسية في أن منصات الإنترنت تستخدم خوارزميات التوصيات بشأن المُحتوى التي تعزز المشكلات ذات الصلة المتمثلة في التطرف وتوطيد المعتقدات داخل المجموعة (Lewis, 2018) وتعمل لأغراض متعددة أو تأتي بنتائج تتعارض والجهود الرامية إلى مواجهة المعلومات المُضلّلة.

وبالنسبة لأدوات الأطراف الثالثة، نظر تقرير «صندوق مارشال الألماني» مؤخراً في 13 شركة ناشئة تهدف إلى استخدام الذكاء الاصطناعي (و/أو التعلم الآلي) لمحاربة المعلومات المُضلّلة. وتشير نتائج هذا التقرير على المستوى الأعلى إلى أنّ «معالجة اللُّغات الطبيعية وحدها لا تستطيع تحديد جميع أشكال التزييف، وأنه من المُرجّح أن تصطدم هذه التكنولوجيا بعدة عقبات قبل أن يتم تنفيذها حتى» (Schiffrin & Goodman, 2019). في الواقع، لا تواجه الأدوات المستقلة القائمة على التعلم الآلي والساعية إلى تحليل الشبكات العقبات المذكورة أعلاه فحسب، بل تواجه عقبات إضافية تخصّ الأدوات القائمة على المنصات، خاصةً إذا كان عليها أن تتفاعل مع بيانات محدودة من وسائل التواصل الاجتماعي ومواقع البحث. وتكمن عقبة إضافية في غياب النفاذ إلى واجهة برمجة تطبيقات API متقاسمة أو بيانات مجمّعة بين شركات الإنترنت. والأمر هذا يشكّل تحدياً بالنسبة إلى الأطراف الثالثة، حيث أنّ المعلومات المُضلّلة لا تبقى على منصة تجارية واحدة بل تنتقل في ما بينها، كما تنتقل عبر المنصات غير التجارية، وتطبيقات المراسلة، والبحث ومشاركة الفيديو. لذلك، يَصْعُب جمع البيانات عن حركة المعلومات المُضلّلة ونشاطها عبر المنصات بصورة فعّالة.

وتشمل نقاط الضعف الأخرى الخاصة بأدوات الأطراف الثالثة مجموعات البيانات المحدودة المتاحة بحُريّة –بسبب قيود الخصوصية والموافقة– والتي تُضاف إلى مجموعات بيانات التدريب الموثوق بها من شركات الإنترنت. علاوةً على ذلك، تتوفّر أدوات الأطراف الثالثة، تماماً كالأدوات القائمة على

[322] https://www.santaclaraprinciples.org/

المنصات، في حركيّة خصمية مع الجهات الفاعلة في حقل التّضليل. ومن المرجح أن تفوّت الحلول الخوارزمية التي تم تدريبها على بيانات سابقة ولم يُعاد تدريبها أو تحديثها أشكالاً جديدة من المعلومات المُضلّلة والمغلوطة، فيكون أداؤها أسوأ بكثير.

حدود أدوات البحث عن الصور، والبحث عن أوجه الشبه، والتحليل الجنائي الإعلامي، وأصل الصور توفّر الصور العكسية والتشابه ذو الصّلة في الصور، عبر استخدام محركات البحث، دقة جيدة عموماً. ومع ذلك، فإنها تعتمد على شمولية الفهرسة التي تقوم بها محركات البحث من أجل تحديد الصور السابقة، ويُسجّل غياب للبحث العكسي للفيديو الذي أثبت فعاليته مع التعديلات الطارئة على أشرطة الفيديو وتحريرها. إذا تمت فهرسة صور زائفة أكثر من الصور الأصلية، فقد يصبح من الصعب استرداد الصورة أو الفيديو الأصلي، خاصة مع مرور الوقت. هذا ويُعتبر البحث العكسي للفيديو معقداً من الناحية الحسابية وهو غير متوفّر حالياً للجمهور على المنصات.

إلى جانب ذلك، تشوب ثغرات تقنية أدوات التحليل الجنائي الإعلامي. فمعظمها لا يعمل بشكل جيد مع المواد الإعلامية المضغوطة ومع مواد إعلامية منخفضة الدقة، أو لا يوفر معلومات سهلة القراءة بالنسبة إلى الإنسان. وإذا ما أشرنا إلى النقص الكبير في فهم التحليل الجنائي الإعلامي في أوساط الصحفيين ووسائل الإعلام والجهات التي تتحقّق من الوقائع، فإنّ أوجه التقدم التي سجّلتها أدوات التحليل الجنائي الإعلامي لا تتماشى دائماً مع احتياجات المجتمع المدني والإعلام (Gregory & French, 2019). وتشمل أوجه القصور هذه معالجة قضايا تتعلّق بنوعية المواد الإعلامية وضغطها، وضرورة اتخاذ قرارات سريعة وشرحها للجمهور المشكّك.

في الختام، لا تزال حواجز سياسية وعملية وأخلاقية هامة تحول دون استخدام أنظمة الذكاء الاصطناعي والتعلم الآلي على نطاق واسع للكشف عن الرسائل والجهات الفاعلة والنشاط، لفائدة علاقتها بحُريّة التّعبير والدقة والتأثير على الفئات الضعيفة من السكّان والشفافية/القدرة على الطعن. وهي غير مناسبة للاستخدام إلا بشكل شبه مميكن وبصفة مساعدة. من جهتها، تُعتبر أدوات تقييم عنصر مُحتوى واحد –مثلاً لتأكيد التلاعب الجنائي التقليدي غير المرتبط بالذكاء الاصطناعي بصورة ما– أكثر قوةً، لكنها تواجه أيضاً مشكلة الثغرات في البيانات وفي قدرة الصحفيين وغيرهم على الاستفادة منها.

على الرغم من التحسينات التي طرأت على العناية الواجبة في مجال حقوق الإنسان بشكل عام ضمن سياسات شركات الإنترنت والبحث والرسائل[323] (Ranking Digital Rights 2019)، لا تزال ثغرات مهمة قائمة (Hogan, 2018). وقد أثارت هذه القضايا انتقادات لعدم الاستثمار المنهجي في تقييمات الأثر التي تشرك بشكل كامل المجتمع المدني وأصحاب المصلحة الآخرين، مع دخول الشركات أسواقاً / مجتمعات جديدة بمنتجات متوفّرة. وبشكل مُماثل، تعرّضت الشركات للانتقاد لعدم تقييمها المخاطر الناشئة في الأسواق القائمة (التقييم اللاحق الذي أجراه فيس بوك لتأثيره في «ميانمار» هو استثناء تم الإعلان عنه رداً على انتقادات المجتمع المدني، Facebook 2018b). ويُسجّل في هذا الصدد نقص في الشفافية يعقّد الرقابة الخارجية على المنصات وخوارزمياتها، بما في ذلك الوصول إلى بيانات تقييم أفضل بشأن التحديد الناجح وتحديد إيجابيات وسلبيات كاذبة. بالإضافة إلى ذلك، تتعرّض الشركات

323 https://rankingdigitalrights.org/index2019/indicators/g4/

للانتقاد لعدم إجرائها «اختبار لإساءة الاستخدام»، حيث «تستثمر المنصات الموارد لمعرفة كيف يمكن إساءة استخدام منصاتها لإلحاق الضرر بالمستخدمين. ويُعتَقد أنّ السياسة الذكية ستحفّز هذا النوع من الاستثمار، تماماً كما رأينا هذا النوع من التحفيز حول الأمن السيبراني في السنوات العشر الماضية» (Soltani, 2018). وبالمثل، هناك استياء من الغياب الظاهر للمقاربات التي تُعنى «بحُريّة التّعبير من حيث التصميم» (Llansó et al., 2020).

تشمل الثغرات في البيانات -كما ذُكر أعلاه- عدم توفير بيانات في الوقت الحقيقي وعبر المنصات وعبر بيانات الشركات للباحثين والصحفيين لإتاحة المجال أمام كشف أفضل. وتتوفّر ثغرات محددة تتعلّق بأدوات المصادقة على المُحتوى الإعلامي السمعي-البصري، بما في ذلك البحث العكسي للفيديو والبحث عن التشابه القوي في المنصات وأدوات المراسلة، فضلاً عن أدوات محسنة لتحديد المصدر توفر إشارات تستطيع الآلة والإنسان قراءتها. بالإضافة إلى ذلك، تبرز الحاجة إلى أدوات أفضل لتحليل «الميمات» memes كمعلومات مُضلّلة (انظر Theisen et al., 2020) وللتمييز بين التهكّم والتّضليل عبر عناصر متعددة من مادة إعلامية.

ومع ازدياد توفّر الصور الزائفة الفجّة والمواد الإعلامية الاصطناعية على نطاق أوسع، تبرز الحاجة إلى الاستناد إلى مجموعات مشتركة من بيانات التدريب (التي يتم تزويدها ومقاربات تزوير جديدة متوفّرة بحُريّة) يمكن تعميمها على أشكال جديدة من التزوير، وإلى أقصى حد ممكن، بالنظر إلى الحركيات الاختراقية، وتكون في متناول مجموعة من المستخدمين مع نتائج قابلة للتفسير (Leibowicz, 2019). وبما أن مؤشرات متعددة ستكون ضرورية عبر مجموعة من التلاعبات، سيتعين على لوحات المعلومات وأدوات الكشف أن تجمع بين اختبارات متعددة تُعنى بالتحليل الجنائي والمصادقة على المُحتوى، وتحويلها إلى أشكال يستطيع الإنسان قراءتها، فتكون مفيدة للصحفيين والمحققين (Verdoliva, 2020). ولابد من استكمال هذه الجهود بالاستثمار في القدرات الجنائية داخل العالمَين الصحفي والتحقيقي لتفسير أشكال جديدة من التلاعب بالصور تستند إلى التعلم الآلي.

8.2.6 توصيات بشأن الاستجابات التقنية والخوارزمية

على ضوء التحديات والفرص المحددة أعلاه، والتبعات الكبيرة للاستجابات الخوارزمية على حُريّة التّعبير، يمكن تقديم التوصيات التالية في مجال السياسة العامة.

توصيات للمنظمات الدولية والدول:

- الاستثمار في رصد تأثير الاستجابات التقنية للتضليل على أُطُر حقوق الإنسان وقياسها وتقييمها.
- دعم تطوير مبادرات مستقلة تتضمّن قياس الأثر وتقييمه لزيادة المعرفة بفعالية الاستجابات التقنية، بما يضمن الشفافية والمعايير القابلة للتحقق.
- العمل مع منصات التواصل عبر الإنترنت لضمان أن تكون الاستجابات التي تباشر بها شفافة وقابلة للقياس بشكل مناسب، والحرص على تنفيذها على نطاق عالمي حقيقي.
- تشجيع الشركات على التعاون بشفافية عبر المعايير الأساسية، وإنتاج بيانات قابلة للمقارنة يمكن استخدامها لوضع لمحة عامة عن المشكلة عبر مختلف الخدمات وأطر السياسات ذات الصلة.

- دعم المبادرات الرامية إلى ضمان الحفاظ على الخصوصية والوصول العادل إلى البيانات الرئيسية من منصات التواصل عبر الإنترنت، لإتاحة المجال أمام البحث والتقييم المستقلّين العالميين في الطريقة التي تؤثر بها الاستجابات الخوارزمية على حدوث التّضليل عبر الإنترنت وانتشاره وتأثيره.
- النظر في إمكانية إقامة هيئات مظالم وطنية مستقلة تمكّن المستخدمين من اللجوء إلى التحكيم المستقلّ في ما يتعلق بالطعون المتعلّقة بعمليات إزالة المُحتوى وتعليق الحسابات الآلية غير العادلة.

توصيات لمنصات التواصل عبر الإنترنت:

- دعم الصناديق التي تُدار بشكل مستقل والمخصّصة للبحث والتقييم المستقلّين لفعالية الاستجابات الخوارزمية التي تطلقها الشركات والمنصات رداً على التّضليل.
- العمل معاً لتحسين قدراتها التكنولوجية على الكشف عن المعلومات المُضلّلة والحدّ منها بشكل أكثر فعاليةً، وتبادل البيانات حول هذا الموضوع، نظراً إلى أن المعلومات المُضلّلة غالباً ما تستغلّ الأساليب العابرة للمنصات.
- الإقرار بحدود الميكنة في مراقبة المُحتوى وتنظيمه، وتوسيع نطاق المراجعة البشرية وعمليات الطعن.
- إصدار تقارير عامة مفصّلة عن الشفافية، بما في ذلك تفاصيل عن عمليات إزالة المعلومات المُضلّلة وتعليق الحسابات التي تنشر معلومات مضللة بشكل مميكن، لأن هذه الاستجابات قد تترك آثاراً هامة على حقوق الإنسان وحُريّة التّعبير.
- إعادة تقييم الأسلوب الذي تسهّل من خلاله تكنولوجيا نماذج الأعمال الحالية جهود أولئك الذين يختلقون المعلومات المُضلّلة وينشرونها (كما هو الحال في الترتيب والتوصيات)، وكيف من الممكن أن يؤدي هذا الأمر إلى تقويض الجهود التقنية الأخرى الرامية إلى التعرّف إلى المعلومات المُضلّلة والتصدي لها.

توصيات لمنظمات المجتمع المدني والباحثين:

- مواصلة رصد النجاحات والمخاطر التي تنطوي عليها الاستجابات التقنية والخوارزمية التي وضعتها منصات التواصل عبر الإنترنت ومواصلة تقييمها بشكل مستقل.
- دراسة الأبعاد التكنولوجية لحملات التّضليل العابرة للمنصات للحصول على منظور أكثر شموليةً عن المشكلة والاستجابات لها.
- العمل على تطوير أدوات جديدة لمساعدة الصحفيين والوكالات الإخبارية وغيرهم من المهنيين في مجال التحقّق في الكشف عن المعلومات المُضلّلة وتحليلها بكفاءة، فضلاً عن الترويج الفعال لكشف زيفها والترويج للمعلومات الموثوقة.
- تعزيز الموثوقية والشفافية في ما يتعلق بأدوارها في الاستجابات التكنولوجية للتصدي للمعلومات المُضلّلة.

3.6 الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات

المؤلفة: كالينا بونتشيفا

تشمل الاستجابات الاقتصادية للتضليل خُطوات ترمي إلى وقف تحقيق الدخل والربح المالي من التّضليل، وتُثني بالتالي عن وضع عناوين مثيرة للانتباه clickbait وإنشاء مواقع إخبارية زائفة وأنواع أخرى من التّضليل الذي يبغي الربح. وتستهدف استجابات حجب الربح المالي أيضاً المُحتوى المضلل أو الكاذب الذي يتم إنشاؤه لأغراض أخرى غير تحقيق الربح وحده، بما في ذلك عندما ينصهر بخطاب الكراهية (في وقت يمكن فيه حجب الربح المالي عن الكراهية القائمة بذاتها). تسعى حملة «وقف الكراهية لتحقيق الربح» StopHateForProfit لعام 2020 إلى حجب الربح المالي عن حزمة «الكراهية والتعصب والعنصرية ومعاداة السامية والتّضليل[324]». ولكن، على الرغم من ذلك، سوف يقوم هذا القسم بمسح هذا النوع من الاستجابات الاقتصادية التي تهدف على وجه التحديد إلى تعطيل عملية تحقيق الربح المستند إلى الإعلانات من المعلومات المُضلّلة على الإنترنت (من خلال مثلاً جعل المواقع الإخبارية الكاذبة غير قابلة للبقاء).

وتجدر الإشارة إلى أن هذا القسم سيغطّي الجوانب الاقتصادية للإعلانات عبر الإنترنت فقط (على أساس كسب المال من المعلومات المُضلّلة عن طريق جذب الإعلانات من خلال أنظمة مميكنة)، وسينظر في المحاولات التي تبذلها شركات الإنترنت لتعطيل هذا الأمر من خلال التدابير الحالية. وينبغي التمييز بين هذه المسألة والدوافع السياسية الأساسية وراء نشر المعلومات المُضلّلة من خلال إعلانات تستهدف الناخبين خلال الانتخابات، والتي تناولها القسم 5.3. وفي الوقت نفسه، ينظر هذا القسم في الاستجابات للجهات الفاعلة التي تسعى مباشرةً إلى تحقيق إيرادات من إعلانات تتضمّن معلومات مضللة. فمن خلال التصدّي لهذه الإعلانات، تثبط منصات التواصل عبر الإنترنت أي عزيمة للاضطلاع بمثل هذا النشاط. وفي هذا الصدد، يُشير مصطلح «حجب الربح المالي» المُستخدم في هذا القسم إلى (1) منع وضع الإعلانات إلى جانب مُحتوى مُضلل، و(2) منع الإعلانات التي تحتوي على معلومات مضللة من الظهور/البقاء على الخدمة التي توفّرها الشركة أو المنصة.

1.3.6 مَن وما الذي تستهدفه استجابات حجب الربح والاستجابات المرتبطة بالإعلانات؟

من خلال المعلومات المُضلّلة، تُوجّه حركة المرور نحو مواقع إلكترونية يمكن فيها استخدام الإعلانات عبر الإنترنت لتحقيق الربح المالي. ويتم تحفيز الحركة هذه من خلال مزيج من المنشورات ذات العناوين المثيرة للانتباه clickbait والمنشورات التي تمّ الترويج لها، أي الإعلانات (التي يمكن أن تكون بطبيعتها عناوين مثيرة للانتباه clickbait). هناك العديد من المواقع الإخبارية الكاذبة وملفات التعريف (على تويتر وفيس بوك مثلاً) والمجموعات الملفقة على الإنترنت التي يتم إنشاؤها في إطار هذه

324 https://www.stophateforprofit.org/

العملية. لإعطاء مثال واحد فقط، قام رجل باستحداث أكثر من 700 ملف تعريف على فيس بوك وبإدارتها بصورة مُعد لها (Silverman, 2017a)، وقام بالترويج لروابط وجذب نقرات إلى مُحتوى مغلوط على مواقع إلكترونية تمكّنت بدورها من زيادة إيرادات من الإعلانات المعروضة إلى جانب المُحتوى هذا (Silverman, 2016). ومن الأمثلة الأخرى على ذلك، استخدام كلّ من Google Adsense وdoubleclick لتمويل شبكة سوافيلوس Suavelos لمواقع المتفوّقين البيض الخادعة في فرنسا (EUDL, 2019c)، بالإضافة إلى شبكة مقرها أفريقيا تضمّ وسائل إعلام غير مرغوب فيها ومواقع تحمل عناوين مثيرة للانتباه تبغي الربح وتنشر معلومات صحية مضللة وتنسخ مباشرةً مقالات صادرة عن وسائل إعلام معينة لجعلها تبدو مشروعة (EUDL, 2020).

تُصمّم المنشورات التي تحمل عناوين مثيرة للانتباه لإثارة استجابة عاطفية في نفوس قرّائها، كالمفاجأة، والإثارة، والتشويق، والفكاهة، والغضب، والتعاطف، والحزن، الأمر الذي يفضي بالتالي إلى تحفيز المشاركة الإضافية عن طريق دفع القراء إلى متابعة الرابط إلى صفحة الويب التي تقوم بدورها بزيادة مشاهدات للإعلانات وإيرادات لمالك الموقع. ويغفل العنوان المثير للانتباه عادةً معلومات رئيسية عن المُحتوى ذي الصلة (Chakraborty et al., 2017) من أجل خلق «فجوة فضول» (Loewenstein, 1994) وإغراء المستخدمين ليضغطوا على الرابط. والأمر هذا، بحكم تعريفه، غالباً ما يعني أن العنوان المثير للانتباه ليس تمثيلاً دقيقاً للمُحتوى الذي يعد به، وأن بإمكانه أن يحتوي على معلومات مضللة كمُحتوى كاذب أو مضلل. وقد تم تشبيه الطبيعة المثيرة والعاطفية للعناوين المثيرة للانتباه على وسائل التواصل الاجتماعي بالصحافة الصفراء، وقد اتّضح أنها تؤمّن «فسحة عامة بديلة للمستخدمين الذين يبتعدون عن الأخبار التقليدية» (Chakraborty et al., 2017). وتبيّن، على سبيل المثال، أن التغريدات المثيرة للانتباه تحتفظ بشعبيتها لفترة أطول، وتجذب مشاركة أكبر، بالمقارنة مع تغريدات غير مثيرة للانتباه (Chakraborty et al., 2017). هذه الخصائص تجعلها ناجحة للغاية في نشر معلومات مضللة ومغلوطة عبر الإنترنت من خلال شبكات المستخدمين الحقيقيين، فضلاً عن استخدامها في العديد من الإعلانات التي تسجّل مشاهدات كبيرة. وقد تكون العناوين المثيرة للانتباه ضمن المُحتوى المباشر أو تأتي على شكل مكوّن في الإعلان.

تُعدّ الإعلانات عبر الإنترنت وسيلة شائعة لتحقيق الدخل من مُحتوى خادع وكاذب على مواقع إخبارية غير مرغوب فيها، حيث يتقاضى واضعوها دخلاً عند عرض الإعلانات إلى جانب المُحتوى غير المرغوب فيه. فعلى سبيل المثال، عندما عُرضت إعلانات (غالباً ما تكون لعلامات تجارية كبرى) على يوتيوب في بداية مقاطع فيديو تحتوي على معلومات صحية مغلوطة، حقّق هذا العرض إيرادات لكل من مالك المنصة (غوغل) وناشر مقاطع الفيديو بشأن علاجات السرطان الزائفة (Carmichael & Gragnani, 2019). وادّعى منتجو المواقع ومقاطع الفيديو الزائفة أنهم يكسبون ما بين 10.000 دولار و30.000 دولار شهرياً من الإعلانات عبر الإنترنت، مثل الرئيس التنفيذي لشركة Disinfomedia (Sydell, 2016).

هناك نوع شديد الفعالية من الإعلانات عبر الإنترنت يُسمى «الإعلانات المظلمة dark ads»، وهي لا تظهر إلا للمستخدمين المستهدفين (على سبيل المثال الناخبين في دائرة هامشية في المملكة المتحدة (Cadwalladr, 2017)) ولا تظهر على الشريط الزمني للمعلن. وقد تم استخدام هذه الإعلانات خلال الحملات السياسية لنشر المعلومات المُضلّلة بهدف التأثير على نتائج الانتخابات (Cadwalladr, 2018). علاوةً على ذلك، ونظراً لطبيعتها الشخصية للغاية، يمكن استخدام الإعلانات المظلمة لاستهداف

مستخدمين محددين بمعلومات مضللة يُرجّح أن يعتقدوا أنها صحيحة. وبما أن الإعلانات المظلمة مخفية عن المستخدمين الآخرين؛ لا يمكن مناقشة المعلومات المُضلّلة الواردة فيها، كما لا يمكن أن ينشر أصدقاء المستخدم أدلة مضادة.

على ذلك، تم استخدام إعلانات فيس بوك، بما في ذلك «الإعلانات المظلمة» الهادفة للغاية مؤخراً لنشر الأكاذيب وبيع منتجات زائفة، عبر استخدام مقاطع فيديو ومواد مسروقة من منصة التمويل الجماعي ذات الشعبية Kickstarter، من بين أمور أخرى (Bitten, 2019). وقد استخدمت عملية احتيالية أخرى بملايين الدولارات على فيس بوك مزيجاً من حسابات فيس بوك المستأجرة، والإعلانات الخادعة، والاشتراكات للاحتيال على مستخدمين أقل دهاءً (هم عادةً من جيل طفرة المواليد– المولودون في الولايات المتحدة الأميركيّة بعد الحرب العالمية الثانية) (Silverman, 2019).

ليست منصات التواصل عبر الإنترنت الأخرى بمأمن من هذه الممارسات. فعلى سبيل المثال، في أواخر عام 2019، نشرت شبكة سوافيلوس Suavelos للمتفوّقين البيض قصة كاذبة مناهضة للمهاجرين على موقع suavelos.eu، كشفتها الجهات التي تتحقّق من الوقائع في وكالة الصحافة الفرنسية AFP [325]. وقد دفع هذا الأمر بمختبر الاتحاد الأوروبي لمُكافحة التّضليل EU DisInfo Lab (EUDL, 2019c) إلى إجراء تحقيق معمّق كشف أن شبكة سوافيلوس Suavelos (التي تتكوّن من عدة مواقع، وصفحات فيس بوك، وقناة يوتيوب، وحسابات تويتر وVKontakte) كانت تكسب المال من الإعلانات عبر Google Adsense أو Doubleclick ومن خلال مُحتوى ذي صلة ومدعوم باستخدام Taboola.

تُصنّف المنشورات المروّج لها على فيس بوك وتويتر كإعلانات ويمكن إعادة نشرها، والإعراب عن الإعجاب بها، والرد عليها.. إلخ. كأي منشور عادي. وتصدر المنصة الاجتماعية الفواتير للمعلنين استناداً إلى مقدار المشاركة التي استحدثتها، أي نقرات الإعجاب والمشاركة والنقرات والمشاهدات.

في كثير من الحالات، يستطيع المعلنون اختيار المستخدمين الذين سيشاهدون المنشور المروّج له، استناداً إلى معلومات مثل الموقع الجغرافي أو النوع الاجتماعي أو الاهتمامات أو نوع الجهاز أو خصائص محددة أخرى. وعندما تستهدف الإعلانات مجموعة ضيقة للغاية من المستخدمين (ما يُسمى بـ «الإعلانات المظلمة»)، مع ملامح محددة للغاية، تُسمى هذه الممارسة الاستهداف الدقيق micro-targeting.

عندما يزور المستخدمون المواقع الإلكترونية ومنصات التواصل الاجتماعي، يُعطون عن طيب خاطر أو عن غير قصد معلومات شخصية لا تُقدّر بثمن، مثل موقعهم، والجهاز المحمول المُستخدم، وعنوان بروتوكول الإنترنت، وتاريخ التصفّح، والوقت الذي يقضونه على مُحتوى معين أثناء التصفُّح، والمشاركة على وسائل التواصل الاجتماعي (مثل: نقرات الإعجاب والمشاركات)، والمزاج (رموز المشاعر emoticons، والصور المتحركة gifs). وتكون ملفات التعريف الاجتماعية إجمالاً غنية بالبيانات وتشمل المزيد من البيانات الشخصية، بما في ذلك عيد الميلاد، وحال العلاقة، وأفراد الأسرة، ومكان العمل، وتاريخ التعليم، وما إلى ذلك. علاوةً على ذلك، يتم تتبّع سلوك المستخدمين على الإنترنت باستمرار

[325] https://twitter.com/AfpFactuel/status/1155125308535840768?s=20

من خلال التكنولوجيا مثل ملفات تعريف الارتباط cookies، وتتبّع البرامج النصية والصور، والإعلانات، ورمز CSS/HTML. وتتيح هذه البيانات كلّها تحديد سمات المستخدمين بصورة مميكنة، ما يفضي إلى الاستهداف الدقيق بالإعلانات و/أو المُحتوى المشخصن.

دعا عدد من صانعي السياسات، لجملة أسباب منها استخدام قوة الاستهداف هذه لنشر الأكاذيب، إلى التحلي بالشفافية وتنظيم الإعلان على الإنترنت باعتبارهما من الخطوات الهامة التي تتيح حجب الربح المالي عن المعلومات المُضلّلة عبر الإنترنت:

> ”يتعيّن على المنصات أن تكيّف سياساتها الإعلانية، بما في ذلك الالتزام بمبدأ «اتبع المال»، مع الحول دون إعطاء حوافز تؤدي الى التّضليل، كالثني عن نشر المعلومات المُضلّلة وتضخيمها بهدف تحقيق الربح. ويجب أن تستند هذه السياسات إلى معايير واضحة وشفّافة وغير تمييزية (Buning et al., 2018)“

2.3.6 مَن الذي تُحاول الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات مساعدته؟

تسعى الاستجابات الهادفة إلى حجب الربح المالي أولاً وقبل كل شيء إلى الحدّ من تداول المعلومات المُضلّلة التي تبغي الربح عبر الإنترنت، وبالتالي حماية المواطنين من المنتجات الاحتيالية، «والعلاجات المعجزة» الضارة، والتّضليل السياسي خلال الانتخابات والاستفتاءات. ولا يتّضح إلى أي مدى يمكن الترويج لمُحتوى آخر أو مُحتوى «مُدرج على قائمة بيضاء» بشكل خاص لأغراض جذب الإعلانات. وهناك مشكلات تطبع الممارسة التي تقضي بالسماح للمعلنين بإدراج المكان الذي يُنشر فيه الإعلان إلى جانب مُحتوى معين على قائمة سوداء (وبالتالي تفادي نشر الإعلان في هذا المكان)- كإدراج أي ارتباط بمُحتوى يتناول فيروس كورونا (سواء أكان صحيحاً أم كاذباً) على قائمة سوداء[326].

ثانياً، تقوم شركات ومنصات الإنترنت جزئياً بتنفيذ تدابير لفحص الإعلانات وضمان شفافيتها من أجل حماية إيراداتها الإعلانية التي تبلغ مليارات الدولارات، نظراً إلى أن الإعلانات باتت تتمّ بشكل متزايد عبر الإنترنت وتصبح مميكنة (WARC, 2019). وقد أدت الشكاوى التي رفعها المستخدمون ومناصرو الحملات إلى سحب المعلنين الرئيسيين الرعاية بسبب نشر إعلاناتهم إلى جانب خطاب الكراهية[327].

يقف افتراض رئيسي وراء الاستجابات الاقتصادية، ومفاده أن شركات الإنترنت والشركات الإعلامية لديها نفوذ كبير يخوّل لها السيطرة على تحقيق الربح المالي من المعلومات المُضلّلة عبر خدماتها وكذا منعه. ثانياً، يُفترض أن نماذج الأعمال الخاصة بالشركات وما يرتبط بها من «اقتصاديات الاهتمام» ليست مواتية جوهرياً للتضليل، وأن قادة هذه المؤسسات على استعداد لاستثمار الوقت والجهد لتنفيذ هكذا استجابات وإنفاذها.

....

326 See https://gfmd.info/press-release-emergency-appeal-for-journalism-and-media-support/

327 https://www.businessinsider.fr/us/facebook-fbrape-ad-boycott-2013-5

ويعتمد التنفيذ الناجح لهذه الاستجابات على المسؤولية الاجتماعية والأخلاقية للشركات وقدرتها على الكشف عن المعلومات المُضلّلة التي تبغي الربح وحجب الربح المالي عنها بشكل فعّال. ونظراً إلى الحجم الهائل للمنشورات والإعلانات المروج لها على خدمات هذه الشركات، تلجأ الاستجابات الاقتصادية في المقام الأول إلى الميكنة الخوارزمية[328]، مفترضةً أنها متطورة بما فيه الكفاية لاكتشاف مسار العمل إزاء التّضليل المرتبط ببعد تحقيق الربح المالي وتحديده. فقط في بعض الحالات، تخضع الإعلانات/ المنشورات المروّج لها والتي تم الإبلاغ عنها للفحص اليدوي. غير أن هذا الفحص ليس دائماً فعالاً[329]. وقد يطرح إشكالية لسببين. أولاً، ينبغي أن تتوفّر إمكانية مناسبة للإنصاف متى أُزيل مُحتوى ما عن طريق الخطأ بموجب هذه الوسائل. ثانياً، من أجل تشغيل عملية مراجعة الإعلانات، ينبغي على المستخدمين الإبلاغ عن الإعلانات أولاً. غير أنه من غير الواضح حالياً ما إذا كان معظم المستخدمين (لا سيّما الأطفال، والبالغين فوق سن الخمسين) يُدركون أن بإمكانهم القيام بذلك. إلى جانب ذلك، يستطيع المستخدمون على بعض المنصات الاطلاع على سبب عرض إعلان معين عليهم والإشارة إلى ما إذا كانوا يرغبون في التوقف عن رؤية إعلانات ينشرها معلن معين. ومع ذلك، تبرز الحاجة إلى مزيد من الأدلة التي تثبت أن المستخدمين يُدركون هذه الإمكانات، حيثما كانت متوفرة، وأنهم بالتالي يستفيدون منها بنشاط.

3.3.6 ما هي المخرجات التي تنشرها الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات؟

دعا تقرير فريق الخبراء رفيع المستوى التابع للاتحاد الأوروبي والمعني بالتّضليل (Buning et al., 2018)، والتقارير الحكومية (مثل (DCMS report, 2018c))، والمنظمات المستقلة التي تتحقّق من الوقائع (مثل (FullFact, 2018))، دعا بقوة، شركات التواصل عبر الإنترنت التي تستضيف الإعلانات إلى جعل جميع البيانات الإعلانية السياسية والمتعلّقة بقضايا معيّنة مدفوعة الثمن متاحة علناً للبحث. ويشمل ذلك معلومات مفصلة عن المؤسسة الإعلانية، وبلد المنشأ، والجهات التي تستهدفها الإعلانات. وقد نوقشت بالفعل تفاصيل التنفيذ الحالي لشفافية الإعلان من قبل شركات التواصل عبر الإنترنت في القسم 6.1 في سياق استجابات تنظيم المُحتوى.

بصورة عامة، تُعدّ مكتبات الشفافية الإعلانية ad transparency libraries عنصراً أساسياً من التدقيق المستقل في الإعلانات السياسية وفي الاستجابات الاقتصادية التي تنفّذها شركات التواصل عبر الإنترنت بهدف الحد من الترويج للمعلومات المُضلّلة وتحقيق الربح المالي منها من خلال الإعلان عبر الإنترنت.

غير أن مداها وفائدتها يبقيان في الوقت الحاضر غير كافيين، ليس فقط من حيث التغطية الجغرافية، ولكن أيضاً من حيث الموضوعات الإعلانية. فعلى سبيل المثال، باستثناء فيس بوك، لا تؤمّن جميع مكتبات الإعلانات الأخرى في الوقت الحالي معلومات الشفافية حول فيروس كورونا والإعلانات ذات الصلة، نظراً إلى أنها ليست واحدة من المشكلات المدرجة في نطاق إعلاناتها. والأمر هذا يعيق بشكل كبير التدقيق المستقل في نطاق إزالة الإعلانات المتعلّقة بجائحة فيروس كورونا.

....

328 https://en-gb.facebook.com/business/help/162606073801742

329 https://www.consumerreports.org/social-media/facebook-approved-ads-with-coronavirus-misinformation/

كما نوقش في القسم 6.1، تتضمّن تقارير الشفافية التي تنشرها منصات التواصل عبر الإنترنت معلومات محدودة حول حجب الربح المالي عن المواقع الإلكترونية والحسابات التي تنشر معلومات مضللة.

4.3.6 من هي الجهات الفاعلة الرئيسية ومن يموّل هذه الاستجابات؟

إن جهود حجب الربح المالي عبارة عن تدابير ذاتية التنظيم تنفّذها منصات التواصل عبر الإنترنت استجابةً لضغوط من الهيئات الإدارية وصانعي السياسات الوطنيين والدوليين. ومن بين الأمثلة على التدابير التنظيمية، والتنظيمية المشتركة الرامية إلى ضمان شفافية حجب الربح المالي والإعلانات عبر الإنترنت؛ يمكن ذكر قانون الإعلانات الصادقة في الولايات المتحدة Honest Ads Act (Warner, 2017) ومدونة الممارسات التابعة للمفوضية الأوروبية (European Commission, 2018c). وتسعى هذه الأخيرة إلى إشراك منصات التواصل عبر الإنترنت (غوغل، وتويتر، وفيس بوك، ومايكروسوفت، وموزيلا-فايرفوكس)، والمعلنين، وقطاع الإعلان. يرد المزيد من التفصيل عن الاستجابات التشريعية وما قبل التشريعية والتنظيمية في القسم 5.1.

نتيجةً لذلك، اتخذت الكثير من منصات التواصل عبر الإنترنت (باستخدام مواردها الخاصة) خُطوات تهدف إلى الثني عن إنتاج المعلومات المُضلّلة لتحقيق مكاسب مالية (بما في ذلك السيطرة على الإعلانات عبر الإنترنت). وعلى غرار الوضع المتّصل بالاستجابات الرامية إلى تنظيم المُحتوى (انظر القسم 6.1)، من الصعب الحصول على أرقام موثوقة عن نفقات المنصات على جهود حجب الربح المالي. وقد وردت في القسم 6.1 السابق نظرة عامة مقارنة عن تدابير حجب الربح المالي وفحص الإعلانات عبر 9 منصات تواصل عبر الإنترنت. وسيرد في ما يلي تحليل للمزيد من التدابير الموجهة نحو الإعلانات بشكل خاص:

- **غوغل**: في نيسان/أبريل 2019 وحده[330]، ذكر غوغل أن ما مجموعه 35,428 معلناً يتّخذون من الاتحاد الأوروبي مقراً لهم انتهكوا سياسة الوصف المضلل من خلال وضع إعلانات مخالفة عبر محرك البحث الخاص بغوغل، ويوتيوب، ومواقع أطراف ثالثة تعرض إعلانات غوغل لأغراض تحقيق الربح. ولكن بما أن سياسات غوغل أوسع من مجرّد التصدي للربح المالي الذي يُحقّق من المعلومات المُضلّلة، فإن الشركة لم تحدد من الناحية الكمية تأثير هذه السياسات على انتشار المعلومات المُضلّلة بشكل محدّد في الوقت الحالي. وخلال الفترة الزمنية نفسها، حيث قام غوغل بالتعرّف إلى 56,968 إعلاناً سياسياً في الاتحاد الأوروبي صادرة عن معلنين تم التحقُّق منهم وقام بتصنيفها وتوفيرها للجمهور، ولكن عند كتابة هذا التقرير لم يكن قد وفّر تقارير شفافية حول الإعلانات المستندة إلى قضايا. وتبرز الحاجة إلى تقرير كمي عن التدابير الآيلة إلى حجب الربح المالي عن مواقع المعلومات المُضلّلة، بما أن دراسة مستقلة حديثة (Global Disinformation Index, 2019) كشفت أن غوغل هو المنصة الإعلانية التي تؤمّن 70% من الإعلانات لمواقع معروفة تنشر معلومات مضللة، مما أدى إلى تحقيق هذه المواقع أكثر من 86 مليون دولار من الإيرادات الإعلانية.

---

[330] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59226

- فيس بوك: خلال الفترة الزمنية نفسها، اتخذ فيس بوك[331] إجراءات ضد 600.000 إعلان في الاتحاد الأوروبي تحتوي على مُحتوى منخفض الجودة أو كاذب أو مضلل كانت قد انتهكت سياساتها. وعلى غِرار غوغل، لا يتّضح كم من هذه الإجراءات كانت على وجه التحديد جهوداً لحجب الربح المالي عن المعلومات المُضلّلة. ويُعتبر فيس بوك فريداً من نوعه في الوقت الحالي من ناحية توفير معلومات عن شفافية الإعلانات السياسية، ولكن أيضاً الإعلانات المستندة إلى قضايا ضمن الفئات التالية: الهجرة، والقيم السياسية، والحقوق المدنية والاجتماعية، والأمن والسياسة الخارجية، والاقتصاد، والبيئة. وقد أتاح هذا الأمر إجراء عمليات تدقيق مستقلة في هذه الإعلانات، بما في ذلك الإعلانات المؤيدة للقاحات وتلك المعارضة لها (Jamison et al., 2019). وفي كانون الثاني/يناير 2020، أُعلن أن المستخدمين سيكونون قادرين على تقليل عدد الإعلانات السياسية والاجتماعية التي يرونها على فيس بوك/انستغرام (Leathern, 2020; Nunez, 2020). ومع ذلك، تُطرح أسئلة رئيسية[332] حول سياسة فيس بوك (Leathern, 2020) المتمثلة في عدم اتباع سياسة غوغل التي تقضي بتقييد الاستهداف عبر الإعلانات السياسية وعدم فحص الإعلانات السياسية بحثاً عن معلومات مضللة، كالسماح بنشر إعلانات سياسية ملفّقة مرتبطة بتغيّر المناخ على المنصة (Kahn, 2019)؛ والترويج لإعلانات تحتوي على بيانات صنّفتها الجهات التي تتحقّق من الوقائع على أنها مغلوطة أثناء الانتخابات في سريلانكا (Wong, 2019b). إنه لمجال يمكن أن يقرّر فيه فيس بوك العودة إلى مجلس الرقابة الجديد[333] الخاص به، نظراً إلى أنّ الصلاحيات المعلنة لهذه الهيئة تقضي بمراجعة قرارات الشركة بشأن إزالة المُحتوى. هناك مجالان آخران بحاجة إلى مزيد من الاهتمام، وهما تأجير حسابات فيس بوك لتحقيق الربح (Silverman, 2019) وحملات جمع التبرعات الفردية وعلى أساس مجموعات صغيرة على فيس بوك، والتي تعمل بنجاح على الترويج لرسائل معارضة للقاح (وغيرها من القضايا الاجتماعية المثيرة للجدل) في انتهاك واضح لسياسات المنصة (Zadrozny, 2019).

- تويتر: بين كانون الثاني/يناير وآذار/مارس 2019 [334]، رفض تويتر 4,590 إعلاناً في الاتحاد الأوروبي لانتهاكها «سياسة ممارسة الأعمال غير المقبولة» الخاصة به، كما رفض 7,533 إعلاناً آخر في الاتحاد الأوروبي لعدم الامتثال «لسياسة الإعلانات ذات الجودة» الخاصة به. ولا يتّضح، مرةً أخرى، كم من هذه كانت معلومات مضللة على وجه التحديد. واعتباراً من تشرين الثاني/نوفمبر 2019، حظر تويتر الإعلانات السياسية على مستوى العالم[335].

- تلقّى يوتيوب 15 مليار دولار من إيرادات الإعلانات في عام 2019 وحده (Statt, 2020). وبشكل عام، يتلقى صانعو مقاطع الفيديو على يوتيوب 55٪ من الإيرادات عند عرض إعلان قبل الفيديو الخاص بهم أو أثناءه، فيما يحتفظ يوتيوب بنسبة 45٪ المتبقية كإيرادات إعلانية (Tameez, 2020). وقد وضعت هذه الخدمة التي يملكها غوغل سياسات[336] حول كيفية تحقيق قنوات يوتيوب الربح

331 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59225

332 https://www.forbes.com/sites/mnunez/2020/01/09/facebook-will-let-you-reduce-the-number-ofpolitical-ads-you-see---but-it-still-wont-stop-politicians-from-lying

333 https://www.nytimes.com/2020/05/06/opinion/facebook-oversight-board.html

334 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59227

335 https://twitter.com/jack/status/1189634360472829952

336 https://support.google.com/youtube/answer/1311392?hl=en-GB

من المُحتوى من خلال كسب إيرادات من عرض الإعلانات. ومتى يتّضح أن مقاطع فيديو وقنوات تنتهك هذه السياسات، يمكن حجب الربح المالي عنها أو إزالتها. وقد أدى هذا الأمر، في بعض الحالات، إلى رقابة ذاتية مارسها صانعو المُحتوى خوفاً من حجب الربح المالي (Alexander, 2020)، بالإضافة إلى اتهامات بالتفاوت في طريقة معاملة مقاطع الفيديو التي يعدّها صانعو المُحتوى من الدرجة الممتازة مقارنةً بتلك التي يضعها صانعو مُحتوى عاديين (Alexander, 2020). وقد أثار مستخدمون حُجب عنهم الربح المالي عن طريق الخطأ على يوتيوب بعض المخاوف حول عدم شفافية قرارات يوتيوب المُتَّخذة، وعدم توفير آلية تظلُّم فعالة، وعدم توفير أي خيارات لاسترداد الدخل الإعلاني المفقود (Goggin & Tenbarge, 2019). بالإضافة إلى ذلك، أظهرت دراسة مستقلة صدرت في كانون الثاني/يناير 2020 أنه على الرغم من سياسات يوتيوب وجهوده المعلنة، كانت الإعلانات المدفوعة من قبل أفضل 100 علامة تجارية تموّل معلومات مضللة عن المناخ (Hern, 2020). ولم تكن العلامات التجارية المتضررة على علم بأن إعلاناتها كانت تظهر قبل عرض مقاطع فيديو تحتوي على معلومات مضللة وأثناء عرضها.

- ريديت Reddit[337]: اعتباراً من 15 أيار/مايو 2020، تقبل الشركة فقط الإعلانات والمعلنين الذين يتّخذون من الولايات المتحدة مقراً لهم، علماً أنها تخضع كلها للمراجعة اليدوية. وفي إطار مقاربة جديدة، ستظهر تعليقات المستخدمين على الإعلانات السياسية لمدة 24 ساعة على الأقل، ويتم تشجيع المعلنين بقوة على التفاعل مع المستخدمين وتعليقاتهم. وهناك أيضاً قسم فرعي جديد لريديت subreddit يتعلّق بشفافية الإعلانات السياسية[338].

- حظر تيك توك[339] الإعلانات المتعلقة بالانتخابات والمناصرة والقضايا من المنصة. ومع ذلك، أثيرت مخاوف (Kozlowska, 2019) من أن شكل تيك توك الموجّه نحو الترفيه، وخوارزميات التوصيات recommender algorithms الاستكشافية، وعدم استعداده النسبي للكشف عن معلومات مضللة واحتوائها تُستغلّ كلّها لنشر رسائل الحملات السياسية، ونظريات المؤامرة، والعلوم الزائفة والترويج لها.

تشتري شبكات الوساطة والتبادل الإعلانيَين المميكنة[340] إعلانات على شبكة الإنترنت آلياً وتبيعها، وقد قُدّرت في عام 2019 بقيمة 84 مليار دولار أميركي أو 65٪ من الإعلانات الرقمية (WARC, 2018). والأسواق الرئيسية المستهدفة بهذا النشاط هي الولايات المتحدة وكندا والمملكة المتحدة والصين والدنمارك (WARC, 2018). ويشمل المُشغّلون الرئيسيون: غوغل[341]، وThe Rubicon Project[342]، وOpenX[343]، وAppNexus[344]،

....

337 https://www.reddit.com/r/announcements/comments/g0s6tn/changes_to_reddits_political_ads_policy/

338 https://www.reddit.com/r/RedditPoliticalAds/

339 https://newsroom.tiktok.com/en-us/understanding-our-policies-around-paid-ads

340 https://digiday.com/media/what-is-an-ad-exchange/

341 https://marketingplatform.google.com

342 https://rubiconproject.com/

343 https://www.openx.com/

344 https://www.appnexus.com/fr

و Criteo [345]. وتشير المصادر المتاحة إلى أنه من بينها كلّها، وحده غوغل التزم حتى الآن بتوفير درجة من الشفافية حول الإعلانات، وفقط في ما يتعلق بالإعلانات السياسية. غير أن هذا الأمر يُعتبر غير كافٍ، لأن مواقع التّضليل والإعلانات المُضلّلة غالباً ما تحقّق الربح من خلال عمليات احتيال اقتصادية بحتةٍ، مثل تجربة المنتجات «المجانية» (Silverman, 2019). وفي الوقت نفسه، خلص تحليل مستقل أُجري في أيلول/سبتمبر 2019 (Global Disinformation Index, 2019) للإعلان البرنامجي على 20,000 نطاق تضليل إلى أنها تحقق ربحاً من دون أي عوائق يتجاوز 235 مليون دولار أمريكي من خلال تبادل الإعلانات. وقد تبيّن أن أعلى حصة في السوق تعود إلى غوغل، الذي جنى أيضاً أعلى مبالغ إيرادات مقدّرة لمواقع التّضليل هذه (أكثر من 86 مليون دولار أمريكي)، يليه AppNexus (أكثر من 59 مليون دولار أمريكي)، و Criteo (أكثر من 53 مليون دولار أمريكي)، وأمازون (أقل بقليل من 9 ملايين دولار أمريكي). فوضع الإعلانات بشكل آلي ومطابقتها مع مُحتوى معين هما من الميزات التي يستطيع منتجو المعلومات المُضلّلة استغلالها بسهولة.

بدأ بعض المُعلنين مؤخراً يمتنعون عن وضع إعلانات على فيس بوك وغوغل وتويتر وغيرها من الخدمات التي تقدمها شركات التواصل عبر الإنترنت، كوسيلة لحجب الربح المالي عن هذه الشركات وتحفيزها على معالجة مسألة المعلومات المُضلّلة بشكل أكثر شمولية وموثوقية، خاصةً في الحالات التي يمكن فيها أن تحرض على العنف أو تقمع التصويت[346]. وقد أسفرت تدابير المقاطعة هذه بالفعل عن خسائر كبيرة[347]. هذا وقد تعاظم هذا الزخم خلال عام 2020 مع حركة «أوقفوا الكراهية لتحقيق الربح» Stop Hate for Profit التي أُدرجت على قائمتها حوالى 600 شركة مشاركة بحلول منتصف عام 2020 [348].

يُعتبر الصحفيون والمجتمع المدني والمنظمات الإعلامية والجهات التي تتحقّق من الوقائع والباحثين جهات فاعلة رئيسية في الكشف عن عمليات الاحتيال على الإنترنت التي تستغلّ المعلومات المُضلّلة أو تستفيد منها؛ وترصد أيضاً تنفيذ الاستجابات الاقتصادية الرامية إلى حجب الربح المالي عن المعلومات المُضلّلة وتُقيّمها وتُسدي النصح بشأنها. وعلى ضوء الطابع الطوعي ذاتيّ التنظيم إلى حد كبير للاستجابات الاقتصادية التي تنفّذها الشركات والمنصات المختلفة، يُعتبر دور هذه الجهات الفاعلة المستقلة أساسياً ومهماً على حد سواء.

5.3.6 دراسة حالة عن الاستجابة: معلومات مضللة حول جائحة فيروس كورونا

في سياق جائحة فيروس كورونا، اتخذت شركات الإنترنت خُطوات لمنع الأشخاص من جني المال من المعلومات المُضلّلة عن جائحة فيروس كورونا، وبالتالي لمحاولة إلغاء الحوافز التي تدفع إلى إنتاج عناوين مثيرة للانتباه ومواقع إخبارية زائفة وأنواع أخرى من المعلومات المُضلّلة التي تهدف إلى الربح من هذا الموضوع.

345 https://www.criteo.com/

346 https://www.bbc.co.uk/news/business-53204072; https://www.bbc.co.uk/news/business-53174260

347 https://www.bloomberg.com/news/articles/2020-06-27/mark-zuckerberg-loses-7-billion-as-companies-drop-facebook-ads

348 https://www.stophateforprofit.org/participating-businesses

وقد سُجّل نوعان رئيسيان من الاستجابات الاقتصادية حتى الآن: حظر الإعلانات وحجب الربح المالي عن المُحتوى الكاذب أو المضلل بشأن جائحة فيروس كورونا.

- على الرغم من أن فيس بوك لا يحظر المعلومات المُضلّلة في الإعلانات السياسية، فقد اتخذت الشركة في هذه الحالة (جنباً إلى جنب مع غوغل[349]) خُطوات استباقية للحد من المعلومات المُضلّلة حول جائحة فيروس كورونا في إعلانات فيس بوك وانستغرام، بالإضافة إلى الحد من الانتفاع الاقتصادي من الجائحة[350]. وجرى ذلك من خلال استبعاد الإعلانات لاختبار الحُزَم، والمطهرات، والكمامات، و«العلاجات» بأسعار مبالغ فيها، والتي غالباً ما يتم الترويج لها من خلال ادعاءات مضللة عن طريق العناوين المثيرة للانتباه. على الرغم من ذلك، ونظراً إلى الأسلوب المُميكن المُستخدم لفحص الإعلانات، وجد المعلنون المأجورون أساليب للالتفاف على الحظر[351] من خلال كلمات مترادفة وسرقة حسابات المستخدمين. وأُحبطت أيضاً الجهود التي بذلها غوغل وبينغ Bing لحجب الربح المالي، ولا تزال تقنية البحث الخاصة بهما تعرض أحياناً صفحات تبيع منتجات مشكوك فيها ترتبط بجائحة فيروس كورونا[352].

- في وقتٍ مُبكر من انتشار الجائحة، فرض غوغل وتويتر أيضاً حظراً شاملاً على جميع الإعلانات التي تذكر جائحة فيروس كورونا، باستثناء تلك التي تضعها الكيانات الحكومية أو مصادر رسمية أخرى معتمدة. وقد نتج عن هذا التدبير تأثير غير مرغوب فيه تمثّل في منع كيانات مشروعة أخرى من إطلاق حملات إعلامية مفيدة من خلال الإعلانات. ونتيجةً لذلك، رفع غوغل الحظر في أوائل نيسان/أبريل 2020 [353]. وظلّ موقف تويتر دون تغيير في أوائل نيسان/أبريل 2020: «يحظر تويتر المُحتوى المروّج له الذي يشير إلى جائحة فيروس كورونا. والاستثناءات الوحيدة لهذا الحظر هي إعلانات الخدمة العامة التي تصدر عن الجهات الحكومية والكيانات فوق الوطنية، والوسائل الإخبارية التي تحمل حالياً شهادة إعفاء للمُحتوى السياسي، وبعض المنظمات التي تربطها شراكة حالية بفريق سياسة تويتر[354]».

- بالإضافة إلى الإعلانات، اتخذ يوتيوب[355] تدابير لضمان تحقيق الربح المالي الأخلاقي من المُحتوى الذي يشير إلى جائحة فيروس كورونا أو يتناوله من خلال طلب وجّهه إلى المؤلفين للتحقق من الوقائع في المُحتوى برمّته واتباع مبادئه التوجيهية. وعندما يتم اكتشاف انتهاكات، تقول الشركة إنها تعمل إما على إزالة المُحتوى المُخالف المُرتبط بجائحة فيروس كورونا، أو الحد من تحقيق الربح منه، أو تعطيل عملية تحقيق الربح المالي مؤقتاً على القناة، على الرغم من أنها لا تقدم إحصاءات حول هذه المسألة.

349 https://blog.google/inside-google/company-announcements/covid-19-how-were-continuing-to-help/
350 https://about.fb.com/news/2020/03/coronavirus/#exploitative-tactics
351 https://www.infosecurity-magazine.com/news/ban-hasnt-stopped-covid19/
352 https://searchengineland.com/a-look-at-googles-recent-covid-19-related-policies-in-search-330992
353 https://www.axios.com/google-coronavirus-advertising-6ff1f504-201c-435a-afe5-d89d741713ac.html
354 https://business.twitter.com/en/resources/crisis-communication-for-brands.html
355 https://support.google.com/youtube/answer/9777243?p=covid19_updates

6.3.6 كيف يتم تقييم الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات؟

أصدرت المفوضية الأوروبية تقييماً مستقلاً لفعالية مدونة الممارسات المتعلقة بالتّضليل (Plasilova et al., 2020) خلص تحديداً إلى ما يلي بالنسبة إلى جهود حجب الربح المالي:

- **فعالية تدابير وضع الإعلان في مكان محدد:** نظراً لغياب بيانات مفصلة بما فيه الكفاية، لم يكن من الممكن تحديد فعالية التدابير التي تنفذها حتى الآن شركات ومنصات التواصل عبر الإنترنت. والاستنتاج هنا هو أنه: «في الوقت الحالي، لا تتمتّع المدونة بمكانة عامة عالية بما فيه الكفاية لممارسة ضغط كافٍ لإدخال تغييرات على المنصات. يجب أن تشير التكرارات المستقبلية للمدونة إلى أن العناوين المثيرة للانتباه هي أداة تُستخدم في التّضليل وفي تحديد مكان عرض الإعلانات».

- **شفافية الإعلانات السياسية وتلك المستندة إلى قضايا معينة:** أقرّ التقييم بالنتائج الإيجابية التي تحققت حتى الآن في هذا المجال، غير أنه أشار إلى وجود مجال كبير للتحسين، لا سيما في ما يتعلق بالإعلانات المستندة إلى قضايا معينة.

- **تمكين مجتمع الأبحاث:** يتسبّب نقص البيانات في مشكلة كبيرة للغاية تعيق البحث المستقل في التّضليل والتقييم الدقيق لفعالية التدابير التي تنفذها شركات التواصل عبر الإنترنت من أجل الحد من انتشار المعلومات المُضلّلة (وفي هذا السياق، نجاح جهود حجب الربح المالي أو فشلها أيضاً).

وخلُص التقييم المستقل للتدابير التي تعتمدها المنصات أيضاً (Plasilova et al., 2020) إلى ما يلي: «يمكن النظر في إنشاء آلية عمل في حال عدم الامتثال لأركان المدونة. ولهذا الغرض، يتعيّن على المفوضية الأوروبية أن تنظر في اقتراحات حول التنظيم المشترك؛ تندرج في إطارها آليات إنفاذ وعقوبات وآليات إنصاف مناسبة». وقد سلّط التقييم الضوء، على وجه الخصوص، على ضرورة ضمان تنفيذ الاستجابات الاقتصادية لحجب الربح المالي رداً على التّضليل بشكل موحّد في جميع الدول الأعضاء في الاتحاد الأوروبي.

على ذلك، اقترح التقييم (Plasilova et al., 2020) عدداً من مؤشرات الأداء الرئيسية التي يتعين تنفيذها. ترد في ما يلي مجموعة مختارة من هذه المؤشرات ذات الصلة مباشرةً بتقييم نجاح الاستجابات الاقتصادية للتضليل:

- **التدقيق في الإعلانات والحدّ من المعلومات المُضلّلة داخلها:** إجمالي حجم الأعمال الذي حصل عليه مشغّلو الإعلانات من الإعلانات المنشورة؛ إجمالي الإيرادات الضائعة (المفقودة) بسبب إغلاق حسابات معينة؛ إجمالي الإيرادات الإعلانية من أهم 100 موقع إلكتروني تم تحديدها على أنها المروّجة الرئيسية للمعلومات المُضلّلة. سيظهر الرصد والإبلاغ المنتظمان لمؤشرات الأداء الرئيسية هذه مع مرور الوقت ما إذا كانت فعالية هذه التدابير تتحسّن.

- **شفافية الإعلانات السياسية والإعلانات المستندة إلى قضايا معينة**: تتضمّن مؤشرات الأداء الرئيسية المقترحة عدد الإعلانات السياسية وتلك المستندة إلى قضايا معينة التي تمّ تصنيفها بشكل خاطىء؛ ونسبة إجمالي حجم الأعمال المحقّق من الإعلانات القائمة على قضايا محددة، إلى جانب الإيرادات المفقودة جرّاء إغلاق الحسابات بسبب خرق سياسات الإعلانات المتعلّقة بقضايا معيّنة.

ومن الشروط الأساسية لقياس مؤشرات الأداء الرئيسية هذه أن تقدم الشركات معلومات أكثر دقةً وشموليةً في مكتباتها الإعلانية مما هي عليه حالياً (Leerssen et al., 2019)، بما في ذلك ضرورة توسيع نطاق وصولها الجغرافي المحدود للغاية حالياً؛ وتجاوز الإعلانات السياسية لضمّ جميع الإعلانات؛ وتحسين توفير المعلومات الهادفة وشفافية المعلن.

على الرغم من الافتقار إلى هذه المعلومات الشاملة، فإن المنظمات الإعلامية والمجتمع المدني والباحثين المستقلين قادرون على إجراء تقييمات وتحقيقات صغيرة النطاق حول دراسات حالة محددة توفر نظرة متعمّقة مهمة حول القيود الحالية للاستجابات الاقتصادية للتضليل. وتشمل الأمثلة ما يلي:

- فيس بوك/انستغرام يتيحان للمعلنين إجراء استهداف دقيق يشمل 78 مليون مستخدم صنّفتهم المنصة على أنهم يهتمون بـ«العلوم الزائفة» (Sankin, 2020)؛
- حالات أُرغمت فيها السُّلطات على اللجوء إلى دعاوى قضائية بسبب عدم تقيّد المنصات بقوانين تمويل الحملات الانتخابية في حال الإعلانات السياسية (Sanders, 2020)؛
- فشل مستمرّ في وقف التضخيم وحجب الربح المالي عن الشبكات المزدهرة على مواقع إخبارية غير مرغوب فيها (EU Disinfo Lab, 2020) أو حسابات تنتهك شروط الخدمة الخاصة بالموقع (Ingram, 2019: Webwire, 2020:, EU Disinfo Lab, 2019c)، على الرغم من الجهود المعاكسة المروّج لها على نطاق واسع؛
- عدم القدرة على التمييز بين الصحافة المشروعة وذات الجودة والمُحتوى الآخر، الأمر الذي يؤدي إلى إجراءات تحجب الربح المالي وتزيل المُحتوى، في انتهاك لحُريّة التّعبير والحق في الوصول إلى المعلومات (Taibbi, 2019)؛
- التقاعس عن الحد من المعلومات والإعلانات السياسية المُضلّلة وتأثيرها السلبي خلال الانتخابات (Reid & Dotto, 2019; Tidy & Schraer, 2019; Who Targets Me, 2019).

7.3.6 التحديات والفرص

تؤمّن هذه الاستجابات الاقتصادية للمعلومات المُضلّلة، إذا نُفّذت على نحو سليم، إمكانية وفرصة واعدة للحد من إنتاج ونشر معلومات مضللة تسعى إلى تحقيق الربح.
ومع ذلك، ما زالت غالبية الاستجابات الاقتصادية في الوقت الحالي بين أيدي جهات فاعلة خاصة تتّخذ قرارات غير متّسقة وغير شفافة. ولا تتوفّر شفافية إعلانية كافية في المعلومات التي تقدّمها شركات التواصل عبر الإنترنت، مما يحول دون التدقيق المستقل من جانب الصحفيين والباحثين. وتُطرح هذه المشكلة بشكل حاد عبر الكثير من المنصات والبلدان، وهي لا تطال فحسب الرعاية الصحية (مثل جائحة فيروس كورونا) أو الإعلانات المتعلقة بقضايا محددة بل أيضاً الإعلانات السياسية.

ويعكس خليط السياسات والمقاربات بين مختلف الشركات التعددية والتنوع، ولكنه يستطيع أيضاً أن يُعيق الاستجابة الفعالة الشاملة للقطاع برمّته لحجب الربح المالي عن المعلومات المُضلّلة. كما يمكن أن يخفي المخاطر المباشرة والدائمة التي تحدق بالحق في حُريّة التّعبير والخصوصية، والتي تتسبّب بها الشركات.

سلّطت جائحة فيروس كورونا الضوء على هذه التحديات، علماً أن هذه الجائحة تمثل أيضاً فرصة مهمة جداً لقيام شركات التواصل عبر الإنترنت باتخاذ إجراءات طارئة لضمان الشفافية الكاملة والمُساءلة ومشاركة أصحاب المصلحة المتعددين. وتستطيع هذه الشركات، بالطريقة هذه، أن تثبت حسن نيتها واهتمامها الصادق بتحسين السياسات والممارسات لدعم المعلومات الموثوقة. وقد ينطوي ذلك على مزيج من سياسات تنظيم المُحتوى لضمان الارتقاء بمستوى الوسائل الإخبارية ذات المصداقية وغيرها من الجهات الموثوقة والمعترف بها؛ التي توفّر المُحتوى، أو لخفض ترتيب المُحتوى الكاذب أو إزالته، وجهود حجب الربح المالي المرتبطة بذلك.

8.3.6 توصيات بشأن الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات

تفضي التحديات والفرص المحددة أعلاه وتبعاتها الهامة على حُريّة التّعبير إلى صياغة توصيات محتملة لاتخاذ إجراءات في هذه الفئة من الاستجابات.

توصيات لشركات ومنصات التواصل عبر الإنترنت:

- تحسين نطاق قواعد بياناتها المتعلقة بشفافية الإعلانات وفائدتها لضمان تغطية جغرافية عالمية؛ وإدراج جميع الموضوعات الإعلانية (وليس الموضوعات السياسية فقط)؛ وتوفير إمكانية الوصول الشامل إلى القراءة الآلية، وهي لمسألة ضرورية لدعم التحاليل الكميّة وتقييم السياسات الإعلانية على نطاق واسع.
- إصدار تقارير مفصّلة وعامة عن الشفافية، بما في ذلك معلومات محددة عن حجب الربح المالي عن المواقع الإلكترونية والحسابات التي تنشر معلومات مضللة.
- فحص الإعلانات السياسية بحثاً عن معلومات مضللة من خلال تسخير الجهود المستقلة القائمة بالفعل في مجال التحقُّق من الوقائع.
- تمكين المستخدمين من التعليق على الإعلانات، مثالياً من لحظة نشرها ولمدة 24 ساعة على الأقل. سيتيح هذا الأمر وسم المُحتوى الضار المحتمل كمقدمة لخطوات أخرى محتملة.
- معالجة مشكلة «تأجير الحسابات» بفعالية (أي الاستخدام المدفوع لحسابات مستخدمين أصلية من قِبل وكلاء المعلومات المُضلّلة) للحد من الممارسة المتمثلة في استغلال حسابات الأفراد لجني الأموال من خلال المعلومات المُضلّلة والإعلانات ذات الصلة.
- العمل معاً على تحسين قدرتها على الكشف عن عمليات تحقيق الربح من المعلومات المُضلّلة والحدّ منها، ذلك أن تحقيق الربح غالباً ما يستغل الأساليب العابرة للمنصات.

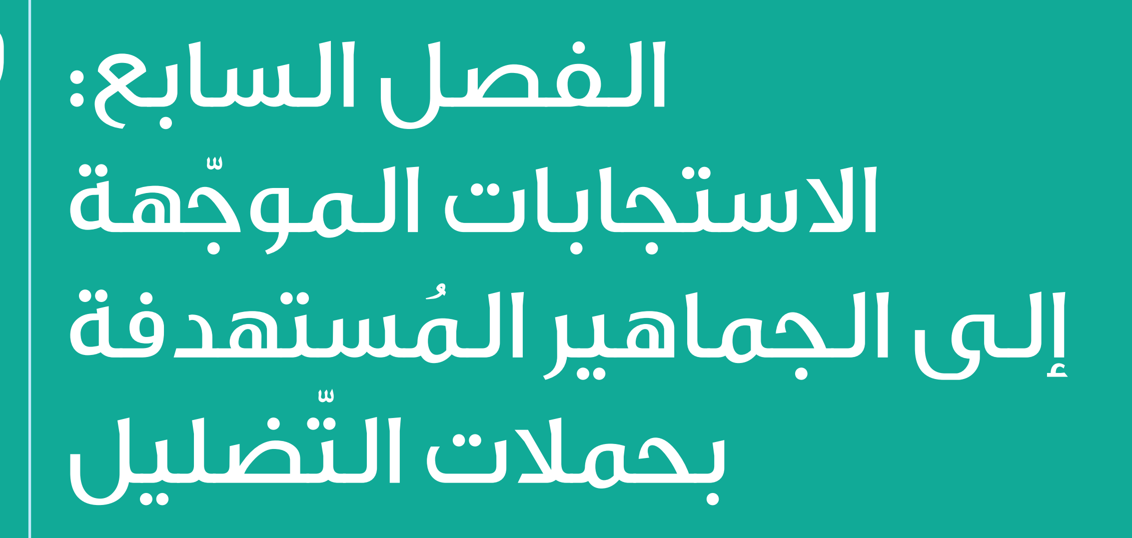
توصيات لشبكات الوساطة والتبادل الإعلانيِّين:

- زيادة رصدها لنطاقات المعلومات المُضلّلة والعمل بالتعاون الوثيق مع الجهات التي تتحقّق من الوقائع والمنظمات المستقلة الأخرى لتنفيذ أساليب فعالة وناجعة وقابلة للتوسّع؛ هدفها حجب الربح المالي عن المواقع التي تنشر معلومات مضللة وعن المُحتوى المضلل.
- تنفيذ تدابير الشفافية الإعلانية التامة، وفقاً لتلك التي توصي بها منصات التواصل عبر الإنترنت.
- العمل معاً من أجل تنفيذ مقاربة مُخطَّط لها؛ تقضي بفحص الإعلانات وضمان الشفافية عبر الشبكات، وهي مقاربة يمكن استخدامها أيضا كوسيلة لتوزيع كُلفة فحص جودة الإعلان وتدابير الشفافية.

توصيات للحكومات والمنظمات الدولية:

- تأمين التمويل المستمر للرصد المستقل وتقييم امتثال جهود حجب الربح المالي التي تنفذها الشركات وشبكات الوساطة والتبادل الإعلانيين.
- التفاوض مع هذه الجهات الفاعلة التجارية بشأن ضمان الشفافية الكاملة والوصول إلى البيانات كشرطين أساسيين للإشراف المستقل على الاستجابات الاقتصادية ذاتية التنظيم للمعلومات المُضلّلة.
- تشجيع شركات التواصل عبر الإنترنت وشبكات تبادل الإعلانات على تنفيذ استجابات مناسبة للتضليل على أساس قواعد القوانين الانتخابية وحُريّة التّعبير، والقيام بذلك في جميع البلدان التي يمكن فيها الوصول إلى خدماتها.
- التشجيع القوي على اعتماد مؤشرات الأداء الرئيسية القابلة للقياس الكمي، والمطالبة بذلك عند الاقتضاء، من أجل إجراء قياس وتقييم مستقلَّين لفعالية الاستجابات القائمة على حجب الربح المالي عن المعلومات المُضلّلة.





الفصل السابع: الاستجابات الموجّهة إلى الجماهير المُستهدفة بحملات التّضليل

1.7 الاستجابات المعيارية والأخلاقية

المؤلفة: جولي بوزيتي

يُناقش هذا الفصل الاستجابات الأخلاقية والمعيارية للتضليل التي تُنفّذ على المستويات الدولية والإقليمية والمحلية. كثيراً ما تنطوي هذه الجهود على إدانة علنية لأعمال التّضليل، أو توصيات و/أو قرارات تتعلق بالاستجابات. وهي تمتد لتشمل المبادرات الرامية إلى ترسيخ القيم والإجراءات على المُستوى الفردي لتساعد في التصدي لانتشار المعلومات المُضلّلة. وبما أنّ الكثير من المعلومات المُضلّلة قد لا تكون غير قانونية (ما لم تُستخدم في الاحتيال المالي أو التحريض على العنف)، يتّسع المجال أمام جهات فاعلة متنوعة لصياغة قرارات أخلاقية تتعلّق بإنتاج المعلومات الملفقة واستضافتها ومشاركتها.

يمكن الغوص في تفاصيل مثلّث القواعد والأخلاقيات والقوانين بطرق مختلفة. يعتبر هذا الفصل أن هذه العناصر قد تكون متوائمة، أو في حالة توتر مع بعضها البعض. وقد تتعارض القواعد والأخلاقيات في بعض الحالات مع الأُطر القانونية، في حين يمكن أن تُفضي الأخلاقيات الشخصية إلى قيام الأفراد بالطعن في قاعدة معينة.

1.1.7 ما هي أهداف الاستجابات الأخلاقية والمعيارية؟

تعمل الاستجابات الأخلاقية والمعيارية للتضليل في المقام الأول على تشكيل القواعد، مما يعزّز الافتراض الاجتماعي المشترك بأن المعلومات المُضلّلة تشكّل تهديداً خطيراً للمجتمع. وتُصمّم الاستجابات للتأثير على القرارات الأخلاقية الفردية الآيلة إلى تحديد إنتاج المعلومات المُضلّلة وتوزيعها وعلى مكافحتهما وردعهما.

لا تتعلّق مجموعة الاستجابات هذه بتوفير حماية «خارجية» لأهداف المعلومات المُضلّلة أو متلقيها، بل تتمحور حول زيادة الجهود لتمكين الناس من بناء مقاومتهم الخاصة للتضليل. وتفترض هذه الاستجابات أنّ سلوكيات المستهدفين تتأثر بالقواعد والأخلاقيات، وأنّ التدخلات ستعزّزها من أجل المُساعدة على «التحصين» ضد المعلومات المُضلّلة والتصدي لها جماعياً.

وتفيد التوقّعات ذات الصلة بأن الناس أخلاقيون وعقلانيون ومنفتحون على فكرة «تحصين» أنفسهم ضد المعلومات المُضلّلة سريعة الانتشار. ومع ذلك، تشير بعض الأدلة إلى أن الكثير من الناس يختارون تصديق المعلومات الخاطئة التي تعزّز معتقداتهم أو تحيزاتهم وتأييدها ونشرها بدلاً من التعامل مع مُحتوى دقيق وذي مصداقية قد يدفع بهم إلى تغيير آرائهم وإضافة فوارق بسيطة إلى هُويّاتهم.

وتماماً كما نوقش في الفصل الثالث (سياق البحث والثغرات)، أكّدت الأبحاث في مجالات علم النفس وعلم الاجتماع على الدور والوظائف المعقّدة للإدراك البشري والمعتقد والأعراف الاجتماعية في النظام البيئي للتضليل. ويتم تسريب الأكاذيب إلى وعي الناس من خلال التركيز على المعتقدات بدلاً من العقل والمنطق، والمشاعر بدلاً من الاستنتاج. ويعتمد انتشار المعلومات المُضلّلة على سياسات التحيز والاستقطاب والحزبية والهوية، فضلاً عن السذاجة والتهكّم وبحث الأفراد عن معانٍ بسيطة

في مواجهة التعقيدات والتغيّرات الكبرى (Posetti & Bontcheva, 2020a). وهذا يفسر لِمَ تشير أبحاث كثيرة إلى صعوبة تغيير المفاهيم الخاطئة، خصوصاً عندما تكتنف الأخطاء الوقائع التي يمكن التعرف عليها، حتى (أو ربما خاصةً) عندما تقوم الجهات التي تتحقّق من الوقائع بكشف المعلومات المغلوطة. علاوةً على ذلك، وكما أثبتت عدة مصادر، يؤدي التكرار والبلاغة إلى تصديق المعلومات غير الدقيقة (Zacharia, 2019). وقد خلُصت أبحاث أخرى إلى أنّ المخاوف الأخلاقية بشأن مشاركة الأكاذيب تتخفض مع التعرّض المتكرر لها (Effron & Raj, 2019). لهذا السبب، ينبغي أن تُراعي الاستجابات الأخلاقية والمعيارية للمعلومات المُضلّلة هذه التعقيدات، وأن تستند إلى بنية واضحة لتتكيّف معها.

تظهر كلمة «ثقة» عشرات المرات في هذا التقرير لأن الكثير من الجهود الرامية إلى الرد على المعلومات المُضلّلة ترتبط بمسألة الثقة –الثقة في الوقائع، والثقة في المؤسسات ذات السمعة الطيبة، والثقة في مصادر المعلومات. ويُنظر إلى الثقة على أنها ناقلة للتضليل وترياق محتمل له على حد سواء– انطلاقاً من مشكلة ما يُسمى «بشبكات الثقة» (تلك الشبكات المؤلفة من العائلة والأصدقاء على وسائل التواصل الاجتماعي) التي تنقل معلومات مضللة (Buchanan & Benson, 2019) وصولاً إلى محاولات مرتبطة بالتّضليل ترمي إلى تقويض الثقة بالصحافة المستقلة النقدية وتشجيع التهكّم (المختلف عن الشكّ) والتفكير التآمري بشأن الأخبار والحقيقة (Ireton & Posetti, 2018; Vaccari & Chadwick, 2020). والثقة عنصر حاسم ولكن بعيد المنال عند التعامل مع المعلومات المُضلّلة. وفي هذا الصدد، تؤثر الاستجابات المعيارية والأخلاقية للتضليل على مسألة الثقة من خلال إنشاء منارة/بوصلة اجتماعية وأخلاقية للسلوك المجتمعي في مجال إنتاج المُحتوى ونقله واستخدامه وتنظيمه.

2.1.7 من وما هي أهداف الاستجابات الأخلاقية والمعيارية؟

تستهدف الاستجابات الواردة في هذه الفئة عادةً، القواعد والأخلاقيات الخاصة بأهداف المعلومات المُضلّلة ومتلقّيها. وتشكّل الدول الأعضاء في المنظمات الحكومية الدولية وصانعو السياسات والجهات الفاعلة القانونية والقضائية محور التركيز الرئيسي لهذه التدخلات، غير أنّ المواطنين على نطاق واسع والمجتمعات على الإنترنت وشركات التواصل عبر الإنترنت وناشري الأخبار والصحفيين مُستهدفون أيضاً.

تعتمد هذه التدخلات على مدى احترام الجهات المستهدفة للمعايير الدولية لحقوق الإنسان (لا سيما حُريّة التّعبير)، وعلى قدرتها على الالتزام بمدونات الأخلاقيات واستعدادها للقيام بذلك، وعلى اهتمامها بتحسين لوائحها التنظيمية وسياساتها ومُمارساتها رداً على تحديات التّضليل.

فعلى سبيل المثال، تعمل المبادرات الموجّهة نحو الصحفيين على افتراض مفاده أن الصحفيين لديهم القدرة والضمير للالتزام بمدونات الأخلاقيات (Storm, 2020) وأنهم مهتمون بتحسين دقة الوقائع في تغطيتهم في مواجهة التحديات المتعلقة بالتّضليل (Taylor, 2020). كما أنها تعتمد إلى حد ما على المعايير والقواعد التي باتت جزءً لا يتجزأ من السياق المهني، واتخذت طابعاً مؤسسياً داخل المؤسسات الإخبارية.

تؤدي الترتيبات المؤسسية، مثل مجالس التنظيم الذاتي، دوراً أساسياً في دعم القواعد والأخلاقيات المرتبطة بكلّ من وسائل الإعلام وشركات التواصل عبر الإنترنت على حد سواء. ويمكن ذكر إحدى المحاولات الأخيرة لتطبيق أُطُر تنظيمية ذاتية أكثر متانةً في هذا المجال، وهي مجلس الرقابة التابع لفيس بوك (Clegg, 2020; Wong, 2020a). إنّه مجلس شبه مستقل يُعين رسمياً ويراجع قرارات إزالة المُحتوى (والأمر لن ينطوي على توجيه القرارات بشأن المُحتوى الذي يتم الاحتفاظ به على الإنترنت في مواجهة الشكاوى). ولا يأتي ميثاق المجلس على ذكر دور المعلومات المُضلّلة ولا المعلومات المغلوطة ولا التحقُّق من الوقائع (Facebook, 2019e)، ولا ما يُسمّيه فيس بوك «السلوك المُفتعل المُعد له» (أي المعلومات المغلوطة والمُضلّلة بصورة منظّمة)، على الرغم من أنها قد تشكّل أسباباً لإزالة المُحتوى. ولا يتّضح ما هي القواعد والمعايير التي ستُطبّق على هذه القرارات إذا كان يُتوقّع من المجلس مراجعتها. من ناحية أخرى، كشف نائب رئيس المجلس المُعيّن حديثاً علناً عن رغبته في «التدقيق» في جهود التحقُّق من الوقائع التي يضطلع بها فيس بوك في إطار عمل مجلس الرقابة، مشيراً إلى وجود «مخاوف جدية» من التحيز السياسي في التحقُّق من الوقائع وتشكيك في التزام الجهات التي تتحقّق من الوقائع بـ«الوقائع» (Allen, 2020).

وترتبط القواعد التي تحدّد انخراط الحكومات في مجال التّضليل بالأُسُس المؤسسية أيضاً، كالبرلمانات والمحاكم والجهات الناظمة المستقلّة للاتصالات.

3.1.7 مَن هي الجهات الفاعلة الرئيسية وما هي الاستجابات التي تُنتجها؟

تتمثّل الجهات الفاعلة الرئيسية التي تطلق الاستجابات المعيارية والأخلاقية للتضليل في ما يلي: المنظمات الحكومية الدولية على الصعيد الدولي (مثل: اليونسكو، ومنظمة الصحة العالمية، وبرنامج الأمم المتحدة الإنمائي) والإقليمي (مثل: الاتحاد الأوروبي، والمجلس الأوروبي، ومنظمة الدول الأمريكية، والاتحاد الأفريقي)؛ وشركات التواصل عبر الإنترنت؛ والوكالات الإخبارية؛ والصحفيين؛ ومنظمات المجتمع المدني. ترد في ما يلي أمثلة محددة عن هذه الاستجابات مع تحليلها.

أ. الاستجابات الحكومية الدولية

على مستوى المنظمات الحكومية الدولية، صدرت عدة توصيات وبيانات وتقارير جديرة بالذكر في إطار محاولة تعزيز القيم والأطر المُصمّمة للتصدي للتضليل ضمن حدود القانون الدولي لحقوق الإنسان.

وفي تطوّر لافت سُجّل في حزيران/يونيو 2020؛ أصدرت أكثر من 130 دولة عضو في الأمم المتحدة ومراقباً رسمياً بياناً بين الأقاليم في سياق جائحة فيروس كورونا. وقد أشار هذا البيان إلى ما يلي: «من الأهمية بمكان أن تتصدى الدول للمعلومات المغلوطة باعتبارها محركاً ضاراً للآثار الثانوية للجائحة التي يمكن أن تزيد من خطر اندلاع النزاع والعنف وانتهاكات حقوق الإنسان والفظائع المرتكبة ضد الإنسانية بصورة جماعية. لهذه الأسباب، ندعو الجميع إلى التوقف فوراً عن نشر المعلومات المغلوطة...». وأشار البيان أيضاً إلى «الدور الرئيسي لوسائل الإعلام الحرة والمستقلة والمسؤولة والتعددية في تعزيز الشفافية والمُساءلة والثقة، وهي أمور أساسية لتوفير الدعم الكافي للجهود الجماعية الرامية إلى الحد من انتشار الفيروس ولامتثال الجمهور لها». وقد أشار البيان، في معرض دعوته البلدان إلى اتخاذ خُطوات لمُكافحة انتشار هذه المعلومات المُضلّلة، إلى ضرورة أن تستند الجهود إلى «حُريّة

التّعبير، وحُريّة الصحافة، وتعزيز أعلى أخلاقيات الصحافة ومعاييرها، وحماية الصحفيين وغيرهم من العاملين في مجال الإعلام، فضلاً عن الترويج للدراية الإعلامية والمعلوماتية، ولثقة الجمهور في العلوم، والوقائع، والإعلام المستقل، والمؤسسات الحكومية والدولية» (UN Africa Renewal, 2020).

الاستجابات على مُستوى الأمم المتحدة:

يُسجّل تدخّل معياري آخر للأمم المتحدة، وهو «الإعلان المشترك بشأن حُريّة التّعبير و«الأخبار الزائفة» والتّضليل الإعلامي والدعاية» لعام 2017 (OSCE, 2017) الذي صدر عن مقرر الأمم المتحدة الخاص المعني بحُريّة الرأي والتّعبير، وممثل منظمة الأمن والتعاون في أوروبا المعني بحُريّة الإعلام، والمقرر الخاص لمنظمة الدول الأمريكية المعني بحُريّة التّعبير، والمقرر الخاص للجنة الإفريقية لحقوق الإنسان والشعوب المعني بحُريّة التّعبير والوصول إلى المعلومات.

وقد جاء هذا الإعلان المشترك، الذي صدر بالتعاون مع منظمتَي المجتمع المدني – Article 19 ومركز القانون والديمقراطية، رداً على سلسلة من التشريعات الصادرة عن دول متعددة تسعى إلى معالجة أزمة المعلومات المُضلّلة من خلال حظر نشر مُحتوى معين وتوزيعه. ويسعى الإعلان إلى معالجة أسباب التّضليل وعواقبه (بما في ذلك الهجمات التي يغذيها التّضليل والتي تشنّها الجهات الفاعلة في الدولة على وسائل الإعلام الإخبارية، والاندفاع نحو تنظيم مُكافحة التّضليل) من خلال إطار القانون الدولي لحقوق الإنسان، مع التأكيد على حقوق حُريّة التّعبير المكرّسة. ويشير الإعلان إلى أن الموقّعين:

> ”إذ يُعربون عن قلقهم من الحالات التي تشوّه فيها السُّلطات العامة سمعة وسائل الإعلام وتخيفها وتهددها، بما في ذلك من خلال التأكيد على أن وسائل الإعلام هي «المعارضة» أو «أنها تكذب» وأن لديها أجندة سياسية خفية؛ وهي حالات تزيد من خطر التهديدات والعنف ضد الصحفيين، وتقوّض ثقة الجمهور بالصحافة بوصفها هيئة مراقبة عامة، وقد تضلّل الجمهور عن طريق طمس الخطوط الفاصلة بين المعلومات المُضلّلة والمنتجات الإعلامية التي تحتوي على وقائع يمكن التحقُّق منها بصورة مستقلة.“

كما أنّ الإعلان، إذ يُدرك أنه بإمكان ما يُسمى «تشريعات الأخبار الزائفة» أن تنتهك حقوق حُريّة التّعبير، لا سيّما من خلال احتواء الصحافة المشروعة وقمعها عن غير قصد (أو عن طريق التصميم)، يشدّد أيضاً على ما يلي:

> ”... لا يقتصر حق الإنسان في نقل المعلومات والأفكار على البيانات «الصحيحة»، فالحق يحمي أيضاً المعلومات والأفكار التي قد تصدم أو تُسيء أو تزعج. وقد ينتهك حظر المعلومات المُضلّلة المعايير الدولية لحقوق الإنسان، في حين أن ذلك لا يبرر في الوقت نفسه نشر بيانات مغلوطة، عن علم أو بشكل متهوّر، من جانب الجهات الفاعلة الرسمية أو التابعة للدولة.“

ويتمثّل الهدف من هذه البيانات في توعية الدول الأعضاء في الأمم المتحدة بمسؤولياتها بموجب القانون الدولي لحقوق الإنسان، وتشجيع الامتثال كوسيلة للثني عن استخدام المعلومات المُضلّلة كأداة لتخويف وسائل الإعلام الإخبارية وغيرها من ناشري المعلومات ذات المصلحة العامة أو تنظيمها باعتبارها وسيلة للحد من حُريّة التّعبير. وتشمل الجماهير المستهدفة بهذه التدخلات أيضاً واضعي السياسات ووسائل الإعلام الإخبارية والجمهور الأوسع نطاقاً.

وتشمل المقاربات ذات الصلة الآيلة إلى تعزيز القيم المعيارية والمعايير الأخلاقية التي اعتمدتها وكالات الأمم المتحدة حملة MILCLICKS# لليونسكو وكتيّب «الصحافة والأخبار الزائفة والتّضليل» (Ireton & Posetti, 2018). وتسعى المبادرة الأولى إلى تعزيز الدراية الإعلامية والمعلوماتية[356]، من خلال التفكير الناقد والإبداع وتعزيز الإلمام والتبادل الثقافي والمواطنة والمعرفة والاستدامة (CLICKS) [357]. وتستهدف المبادرة جمهور الشباب، وقد صُمّمت لتعزيز الانخراط العملي في مجال المعلومات عبر الإنترنت – سيّما وأنه حجر الزاوية في الاستجابات المتوسطة وطويلة الأجل للتضليل– ولتعزيز مفهوم «التفكير المُتأنّي قبل مشاركة المُحتوى» ThinkBeforeSharing#. ويتم، في هذا الصدد، التشجيع على ممارسة معيارية تتمثل في المساءلة في حقل الاتصالات، والتفكير المستنير والأخلاقي في كيفية تعامل الأفراد مع المُحتوى. من جهته، يهدف دليل اليونسكو إلى دمج المقاربات الأخلاقية والنقدية والمتعلّقة بالمساءلة ضمن التثقيف والتدريب الصحفيين لمُكافحة التّضليل. ويعتمد الدليل إطاراً أخلاقياً للدفاع عن الصحافة في مواجهة المعلومات المُضلّلة: «تُعدّ الصحافة الأخلاقية التي تقدّر الممارسات الشفافة والمساءلة جزءً مهماً من ترسانة الأسلحة في معركة الدفاع عن الوقائع والحقيقة في عصر «اضطراب المعلومات» (Ireton & Posetti, 2018). ويُشير الدليل أيضاً إلى ما يلي:

> ”تُعدّ المعايير المهنية للصحافة المحكومة بالأخلاقيات والمساءلة خطّ دفاع بالغ الأهمية ضد التّضليل والمعلومات المغلوطة. وقد تطورت القواعد والقيم التي توفّر الإرشاد للأشخاص الذين يمارسون الصحافة على مرّ السنين لمنح الصحافة مهمتها وطريقة عملها المتفردة. في المقابل، تدعم هذه القواعد والقيم المعلومات القابلة للتحقّق والتعليق المستنير الذي يتم مشاركته جماعياً خدمةً للمصلحة العامة. وهذه هي تحديداً العوامل التي تعزّز مصداقية الصحافة، وبالتالي فقد أُدمجت في متن هذا المقرّر». (Ireton & Posetti, 2018).“

الاستجابات على الصعيد الإقليمي:

شكّلت المبادرات السياسية، ومواثيق الالتزامات، والتحقيقات والأبحاث الهادفة التي اضطلعت بها المفوضية الأوروبية ومجلس أوروبا مساهمةً قيّمة اندرجت ضمن محاولة شاملة لتعزيز الاستجابات المعيارية والأخلاقية للتضليل في أوروبا.

والمفوضية الأوروبية تروّج لفهم معياري مفاده أن المعلومات المُضلّلة يمكن أن «تسبّب ضرراً عاماً، وأن تشكل تهديداً للعمليات السياسية ولعمليات صُنع السياسات الديمقراطية، بل قد تُعرّض صحة

356 https://en.unesco.org/themes/media-and-information-literacy

357 https://en.unesco.org/MILCLICKS

مواطنـي الاتحـاد الأوروبـي وأمنهـم وبيئتهـم للخطـر» (European Commission, 2019). وتحـدّد المفوضيـة المقاربـة التـي تتبعهـا ونيّتهـا فـي مجـال السياسـة العامـة الراميـة إلـى مُكافحـة المعلومـات المُضلّـة فـي مخـزن السياسـات الخـاص بهـا علـى شـبكة الإنترنـت. ويمكـن تلخيـص الأهـداف كالتالـي:

> "يشـكّل تعـرّض المواطنيـن للتضليـل واسـع النطـاق، بمـا فـي ذلـك المعلومـات المُضلّـة أو المغلوطـة، تحديـاً رئيسـياً بالنسـبة إلـى أوروبـا. وتعمـل المفوضيـة علـى تنفيـذ مجموعـة واضحـة وشـاملة وواسـعة مـن الإجـراءات لمعالجـة انتشـار المعلومـات المُضلّـة عبـر الإنترنـت وتأثيرهـا فـي أوروبـا ولضمـان حمايـة القيـم والنظـم الديمقراطيـة الأوروبيـة (European Commission, 2019)."

وقـد استرشـدت هـذه المقاربـة بالمنـح الدراسـية التعاونيـة ومشـاورات الخبـراء، بمـا فـي ذلـك عمـل فريـق الخبـراء رفيـع المسـتوى التابـع للاتحـاد الأوروبـي حـول «الأخبـار الزائفـة» والتّضليـل عبـر الإنترنـت. وقـد رفـع الفريـق، فـي تقريـره النهائـي (Buning et al., 2018)، سلسـلة مـن التوصيـات التـي تؤكـد علـى قيـم الخصوصيـة والأخلاقيـات المهنيـة والمسـؤولية الاجتماعيـة.

تتمثّـل إحـدى المبـادرات الناجمـة عـن مقاربـة السياسـة المعياريـة التـي اعتمدتهـا المفوضيـة الأوروبيـة (European Commission, 2018a) فـي خطـة العمـل الآيلـة إلـى مُكافحـة التّضليـل (European Commission, 2018a)، وهـي خطـة صُمّمـت للتعامـل مـع الأفعـال القانونيـة للتضليـل، وصيغـت للتصـدي للتهديـدات الجيوسياسـية ولضـرورة تعزيـز القيـم الديمقراطيـة الأوروبيـة: «جـاءت خطـة العمـل هـذه اسـتجابةً للدعـوة التـي وجههـا مجلـس أوروبـا فـي عـام 2019 لاتخـاذ تدابيـر «لحمايـة الأنظمـة الديمقراطيـة للاتحـاد ومُكافحـة المعلومـات المُضلّـة، بمـا فـي ذلـك فـي سـياق الانتخابـات الأوروبيـة المقبلـة».

ومـن النتائـج العمليـة الأخـرى التـي تُركّـز علـى الأخلاقيـات، يمكـن ذكـر مدونـة الممارسـات المتعلقـة بالتّضليـل (European Commission, 2018c) التـي وضعتهـا المفوضيـة الأوروبيـة والتـي نُشـرت فـي أواخـر عـام 2018 مـع التأكيـد علـى مـا يلـي: «هـذه هـي المـرة الأولـى التـي يوافـق فيهـا القطـاع، فـي جميـع أنحـاء العالـم وعلـى أسـاس طوعـي، علـى معاييـر التنظيـم الذاتـي لمُكافحـة المعلومـات المُضلّـة». ويشـمل الموقّعـون فـي الوقـت الحالـي فيـس بـوك وغوغـل وتويتـر وموزيـلا فايـر-فوكـس ومايكروسـوفت، إلـى جانـب ثمانـي جمعيـات تجاريـة إعلانيـة (European Commission, 2018d). وتتضمّـن الأهـداف المُعلنـة للمدونـة الشـفافية فـي الإعلانـات السياسـية، وإن لـم تـرد فيهـا أي إشـارة إلـى الدقـة أو التحقُّـق مـن الوقائـع فـي حقـل الإعلانـات السياسـية. والأمـر هـذا ذو صلـة بالمناقشـات الجاريـة والمرتبطـة بالإعلانـات السياسـية خـلال الانتخابـات، وهـي مناقشـات شـهدت دعـوات لاعتمـاد «الحقيقـة ضمـن معاييـر الإعـلان السياسـي»[358].

طلـب مجلـس أوروبـا إعـداد تقريـر بحثـي تأسيسـي يشـدّد علـى دور الأخلاقيـات والقواعـد المهنيـة فـي مُكافحـة مـا سـمّاه «اضطـراب المعلومـات» (Wardle & Derakhshan, 2017). ويقـدم التقريـر مجموعـة مـن التوصيـات للـدول وشـركات التكنولوجيـا ووسـائل الإعـلام الإخباريـة والمجتمـع المدنـي والمموليـن. فبالإضافـة إلـى تعطيـل الانتخابـات الديمقراطيـة، حـدد التقريـر الشـاغل الأكبـر الـذي يتطلـب الاهتمـام، وهـو «... الآثـار

---

[358] انظر النقاش أدناه.

طويلـة الأجـل لحمـلات التّضليـل المُصمّمـة خصيصـاً لـزرع الشـكّ والارتبـاك ولزيـادة حـدّة الانقسـامات الاجتماعية–الثقافيـة القائمـة باسـتخدام التوتـرات القوميـة والإثنيـة والعِرقيـة والدينيـة» (الصفحـة 4). ويشـير ذلـك إلـى ضـرورة اعتمـاد اسـتجابات للتضليل تعتـرف بالمخاطـر الناجمـة عـن تقاطعـه مـع خطـاب الكراهيـة، وتسـعى إلـى تعزيـز القواعـد والقيـم مثـل: المسـاواة العِرقيّـة والجنسـانية والتسـامح الدينـي.

وقّعـت منظمـات حكوميـة إقليميـة تمثّـل أمريـكا اللاتينيـة وأفريقيـا، إلـى جانـب منظمـة الأمـن والتعـاون فـي أوروبـا، علـى «الإعـلان المشـترك بشـأن حُريّـة التّعبيـر و«الأخبـار الزائفـة» والتّضليـل الإعلامـي والدعايـة» لعـام 2017، المُشـار إليـه أعـلاه. وأكّـد ممثـل منظمـة الأمـن والتعـاون فـي أوروبـا المعنـي بحُريّـة الإعـلام مـا يلـي: « ... فـي جميـع الأوقـات، لا سـيّما فـي الأوقـات الصعبـة، لا يمكـن اعتبـار حجـب وسـائل الإعـلام أو حظرهـا رداً علـى ظاهـرة التّضليـل والدعايـة، لأنهمـا يؤديـان إلـى أعمـال تعسـفية ذات دوافـع سياسـية. إنّ القيـود المفروضـة علـى حُريّـة الإعـلام مـن أجـل المنفعـة السياسـية تـؤدي إلـى الرقابـة، وعندمـا تبـدأ الرقابـة، لا تتوقـف أبـداً. بـدلاً مـن ذلـك، يكمـن الجـواب فـي المزيـد مـن النقاشـات وفـي تعدديـة وسـائل الإعـلام[359]». بالإضافـة إلـى ذلـك، دعمـت منظمـة الأمـن والتعـاون فـي أوروبـا ورش عمـل قُطريّـة مخصّصـة لترسـيخ قواعـد حُريّـة التّعبيـر فـي الاسـتجابات للمعلومـات المُضلّلـة، مـع تجهيـز الـدول الأعضـاء عمليـاً للاسـتجابة للمعلومـات المُضلّلـة (OSCE, 2017).

ب. استجابات المجتمع المدني:

تشـمل اسـتجابات كثيـرة للتضليل اعتمدهـا المجتمـع المدنـي مبـادرات تسـعى إلـى تعزيـز القيـم الديمقراطيـة وأُطُـر حقـوق الإنسـان التـي تدعـم قواعـد مثـل حُريّـة التّعبيـر، والنفـاذ إلـى المعلومـات، واحتـرام الخصوصيـة، والمسـاواة بيـن الجنسَـين وبيـن مُختلفـي الأعـراق. ويـرد فـي القسـم «د» أدنـاه تفاصيـل حـول عـدد مـن هـذه التدخـلات، التـي تُعنـى بالتقاطـع بيـن المعلومـات المُضلّلـة وخطـاب الكراهيـة.

إلـى جانـب ذلـك، صُمّمـت مبـادرات متعـددة تُعنـى بالدرايـة الإعلاميـة والمعلوماتيـة أطلقتهـا منظمـات المجتمـع المدنـي ويـرد وصفهـا فـي الفصـل التالـي عبـر إدراج مكوّنـات معياريـة وأخلاقيـة قويـة فـي جوهرهـا. وتسـعى هـذه المبـادرات إلـى تحفيـز الاسـتجابات الأخلاقيـة للتضليـل علـى مسـتوى القاعـدة الشـعبية.

مـن المبـادرات البـارزة التـي أطلقهـا المجتمـع المدنـي مـن أجـل معالجـة تلـوث المعلومـات، منتـدى منظمـة «مراسـلون بـلا حـدود» حـول المعلومـات والديمقراطيـة[360] الـذي يسـتند إلـى إعـلان دولـي أقـره 38 بلـداً. وتقـوم هـذه المبـادرة بتقييـم قواعـد شـبكات التواصـل العالميـة وهيكليتهـا، والتحقيـق فـي إجـراءات الشـركات، ورفـع توصيـات، وتسـهيل التنظيـم والتنظيـم الذاتـي، وطلـب إجـراء أبحـاث ودعـم الصحافـة.

359 https://www.osce.org/fom/319286

360 /https://informationdemocracy.org، للعلـم: مؤلفـة هـذا الفصـل هـي عضـو فـي اللجنـة التوجيهيـة لفريـق العمـل المعنـي بأزمـة المعلومـات التابـع للمنتـدى: /https://informationdemocracy.org/working-groups/concrete-solutions-against-the-infodemic

ج. استجابات شركات ومنصات التواصل عبر الإنترنت والصناعات الإخبارية:

يمكن إعداد قائمة بمجموعة من الاستجابات المعيارية والأخلاقية للتضليل أطلقتها شركات ومنصات التواصل عبر الإنترنت والمؤسسات الإخبارية.

قرّر تويتر حظر إعلانات المرشحين السياسيين من موقعه قبل انتخابات المملكة المتحدة في عام 2019، وأعلن الرئيس التنفيذي للشركة ومؤسّسها جاك دورسي عبر تغريدة: «لقد اتخذنا قراراً بوقف جميع الإعلانات السياسية على تويتر على مستوى العالم. نحن نعتقد أنه ينبغي أن يستحقّ المرء النطاق الذي تبلغه الرسائل السياسية، لا أن يشتريه» (Dorsey, 2019). وحذا غوغل حذوه بعد أسابيع قليلة، مكرّراً التزام تويتر بمنع الاستهداف الدقيق للمستخدمين بإعلانات ذات طابع سياسي. بالإضافة إلى ذلك، وعد غوغل بحظر «الصور الزائفة الفجّة» وما أسماه «الادعاءات الكاذبة بشكل واضح» لمحاولة حماية نزاهة الانتخابات وتعزيز الثقة في العمليات الديمقراطية (Wong, 2019a).

نتيجةً لذلك، تعرّض فيس بوك لضغوط أخلاقية متزايدة لمعالجة سياساته المتعلقة بالمعلومات المُضلّلة والمغلوطة المتصلة بالإعلانات السياسية والخطابات على موقعه (انظر القسم 4.1) – والتي تمّ التركيز عليها بصورة حادة عقب فضيحة Cambridge Analytica [361] – بعد أن قرّر عدم تطبيق معايير التحقُّق من الوقائع على أنواع معينة من الإعلانات السياسية (Eisenstat, 2019; Stewart, 2019). وقد نظر فيس بوك في إمكانية تقييد الاستهداف الدقيق للمستخدمين من قبل الجهات الفاعلة السياسية (Glazer, 2019). ومع ذلك، أعلنت الشركة في نهاية المطاف أنها لن تحدّ من مثل هذا الاستهداف الدقيق، وأنها لن تتّخذ أي إجراء لمنع السياسيين من الإدلاء بادعاءات مغلوطة في منشوراتهم، ولا في الإعلانات المدفوعة، قبل الانتخابات الأميركية لعام 2020 (Romm et al., 2020). وبموجب هذه السياسة (انظر القسم 4.1)، استبعدت الشركة أنواعاً معينة من محتويات الإعلان السياسي من أعمال التحقُّق من الوقائع التي تعاقدت بشأنها مع أطراف خارجية (مما يعني أنها لا تصنّف هذا النوع من المُحتوى على أنه زائف ومضلل) (Hern 2019a; Van Den Berg & Snelderwaard, 2019). غير أن الشركة طبّقت بروتوكولات جديدة في الولايات المتحدة، إذ كان بإمكانها أن تطلب من شركائها المعنيين بالتحقُّق من الوقائع تقييم صدق الإعلانات غير السياسية على فيس بوك (Hern, 2019b).

وقدّم فيس بوك حجة معيارية مفادها أنه من غير المناسب، بشكل عام، لشركة خاصة أن تكون حكماً للحقيقة في حال الإعلانات السياسية (Gilbert, 2019). وفي تدوينة نُشرت في عام 2019، قال نائب رئيس فيس بوك للشؤون العالمية والاتصالات «نيك كليغ» إنّ حُريّة التّعبير هي «مبدأ أساسي مطلق بالنسبة لفيس بوك» (Clegg, 2019). ووفق ما أشار إليه مقرر الأمم المتحدة المعني بحُريّة الرأي والتّعبير «دايفيد كاي»، ينبغي ألا يحول تجنب تأدية دور حكم الحقيقة دون قيام فيس بوك باتخاذ إجراءات ضد الأكاذيب الواضحة (Kaye, 2020b). ويقتضي النقاش المعياري في الممارسة العملية بإقامة توازن بين تفسير الشركة لحُريّة التّعبير والقيود الفعلية على التّعبير المنصوص عليها في معايير مجتمع فيس بوك، وكيف يمكن مقارنة هذه القيود مع تلك التي يُسمح للدول بفرضها بموجب القانون الدولي لحقوق الإنسان. والنتيجة كانت إثارة جدل حول ما إذا كانت الحالات تنتهك معايير مجتمع فيس بوك أو تثير قضايا التقييد بموجب المعايير الدولية (التي من المتوقع أن يحترمها القطاع الخاص، وفقاً لمبادئ

[361] https://www.theguardian.com/news/series/cambridge-analytica-files

الأمم المتحدة المتفق عليها في «تقرير روجي» Ruggie[362]). ومن الأمثلة على ذلك نظريات المؤامرة، التي يتم التسامح معها من حيث المبدأ على الخدمات، ما لم يُعتبر أنها تحتوي على مُحتوى كاذب أو مضلل يمكن أن يسبّب ضرراً وشيكاً. على الرغم من ذلك، يتوفّر دليل عن مقاربات أكثر تقييداً برزت في منتصف عام 2020، عندما أزال فيس بوك حوالى 800 صفحة ومجموعة، وقيّد حوالى 2000 حساب انستغرام ترتبط بنظرية المؤامرة QAnon (Facebook, 2020b).

دفعت المخاوف الأخلاقية بشأن مقاربة فيس بوك تجاه التحقُّق من الوقائع في الإعلانات السياسية بمئات الموظفين في الشركة إلى التوجّه برسالة إلى الإدارة كتبوا فيها ما يلي: «إنّ حُريّة التّعبير والخطابات المدفوعة ليسا الأمر نفسه». وادّعوا أن سياسات تجنّب التحقُّق من الوقائع في الإعلانات الصادرة عن السياسيين والأحزاب السياسية والمنتسبين إليها «تشكّل تهديداً لما يرمز إليه فيس بوك». وذكروا أن السياسة لا تحمي الأصوات، ولكنها بدلاً من ذلك «تتيح للسياسيين استخدام منصتنا كسلاح لاستهداف الأشخاص الذين يعتقدون أن المُحتوى الذي نشرته شخصيات سياسية جدير بالثقة» (New York Times, 2019).

لابدّ من الإشارة مع ذلك إلى أن سياسة فيس بوك لا تزال تسمح برفض الخطابات المباشرة أو الإعلانات من جانب المرشحين السياسيين وشاغلي المناصب الحاليين والأحزاب السياسية والمنتسبين إليها إذا كان الأمر يشكل تهديداً مباشراً للسلامة «في العالم الحقيقي»، أو إذا كان يتعارض مع سياسات الشركة المتعلّقة بقمع الناخبين (Facebook, 2019d). فعلى سبيل المثال، في 30 آذار/مارس 2020، أزال فيس بوك وانستغرام مقاطع فيديو للرئيس البرازيلي «جاير بولسونارو» بحجة نشر معلومات مضللة عن فيروس كورونا وبالتالي انتهاك شروط استخدام المنصتَين. فالشروط هذه لا تسمح «بنشر معلومات مغلوطة» قد تسبّب ضرراً جسدياً للأفراد، كما قال فيس بوك (BBC News, 2020b). غير أن هذه المعايير لا تُطبق على الصعيد الدولي بشكل موحّد. فعلى سبيل المثال، لم تُحذف المنشورات التي نقلت حديثاً للرئيس الأمريكي السابق «دونالد ترامب» يتكهّن فيه أن مواد التبييض قد تكون علاجاً محتملاً لفيروس كورونا (Suárez, 2020).

على الرغم من أن الرئيس التنفيذي لشركة فيس بوك «مارك زوكربيرج» ذكر أن الترويج لمواد التبييض كعلاج لفيروس كورونا كان من فئة «المعلومات المغلوطة» التي ستتم إزالتها على الفور –بسبب «الخطر الوشيك» –، فقد اعتبرت الشركة أن بيان ترامب لم ينتهك السياسة لأنه لم يتوجّه فيه على وجه التحديد إلى الناس ليتناولوا مواد التبييض. ومنذ ذلك الحين، أزال فيس بوك مقطع فيديو زعم فيه الرئيس الأمريكي السابق أن الأطفال «مُحصّنين بشكل افتراضي» ضد فيروس كورونا (Kang and Frenkel 2020). وقد اعترف كبير مدراء سياسة فيس بوك السابق «ريتشارد آلان» بأن فيس بوك يطبّق معاييره بشكل مختلف حول العالم، شارحاً الاختلافات في المعاملة من حيث «الخطر» الناتج عن قرب دولة ما من الولايات المتحدة وحجمها (Suárez, 2020). وفي أيلول/سبتمبر 2020، نشر موقع Buzzfeed مقتطفات من مذكرة وضعتها عالمة بيانات سابقة في فيس بوك ادّعت فيها أن الشركة قد تخلّت بانتظام عن مسؤوليتها في التعامل مع معلومات سياسية مضللة يُحتمل أن تتسبّب بضرر جسدي خارج الدول الغربية. وقد أشارت إلى أمثلة من بلدان نامية متعددة (Silverman, Mac and Dixit, 2020)

[362] https://www.business-humanrights.org

على ذلك، أُثيرت مخاوف ذات صلة في تقرير لمجلس اللُّوردات البريطاني اعتبر أن «فيس بوك قد أعاق بشكل متعمد مبادرة التحقُّق من الوقائع من جانب أطراف ثالثة من خلال إعفاء جميع السياسيين المنتخبين والمرشحين للمناصب من أي تحقّق من الوقائع» (House of Lords, 2020).

لتحديد ما إذا كان المُحتوى رأياً أو واقعاً –إن كان هذان الأمران متشابكين بشكل وثيق في بند معين–، ينبغي إصدار حكم أخلاقي، من بين جملة أمور. وكما نوقش في القسم 4.1، يُسلّط هذا الأمر الضوء على الثغرات السياسية التي لا يجوز في إطارها تصنيف المعلومات المُضلّلة على هذا النحو، أو على الحالات التي تقوم فيها الشركة بإزالة تصنيفات التحقُّق من الوقائع التي تشير إلى الزيف؛ لأنّ الأكاذيب تختلط بالرأي (الذي تعتبره سياسة فيس بوك مُعفَى إلى حد كبير من التحقُّق من الوقائع)، مما يؤدي إلى أحجية بشأن ما يشكّل استجابة مناسبة على المستوى الأخلاقي. فعلى سبيل المثال، قام فيس بوك بإزالة التصنيفات الناتجة عن التحقُّق من الوقائع التي وضعتها جهات خارجية معنية بالتحقُّق من الوقائع على مُحتوى اعتُبر مجرّد رأي (Penney, 2020; Pasternack, 2020; Grossman and Schickler, 2019). وهناك أيضاً تقارير عن ضغوط تمارسها الشركة على الجهات الخارجية التي تتحقّق من الوقائع، في إشارة إلى التقييم الذي تضعه هذه الجهات بشأن المُحتوى الذي يشكّل «رأياً» و«مناصرة»، واتهام الجهات التي تتحقّق من الوقائع عن طريق الخطأ بالتحيز لفائدة تصنيف المعلومات العلمية المُضلّلة، في ظل وجود شفافية محدودة للغاية (Pasternak, 2020).

في حين يرتبط فيس بوك بشراكات رسمية طويلة الأمد في مجال التحقُّق من الوقائع مع الكثير من الوكالات الإخبارية والمنظمات غير الحكومية[363] ذات السمعة الطيبة، التي ذكر عدد منها توفّر حافز مدفوع بالمهمّة للمشاركة (Funke & Mantzarlis, 2018a)، أثارت المبادرة انتقادات أخلاقية من بعض الصحفيين. وكان من بين النقاد أولئك الذين يشاركون بنشاط بمهام التحقُّق من الوقائع ويشعرون أن التعاون تضارب مع المعايير المهنية (Levin, 2018). وقد انسحب عدد من الشركاء المعنيين بالتحقُّق من الوقائع في نهاية المطاف من الترتيب، في خضمّ نقاشات حول الأخلاقيات المهنية المرتبطة بتشغيل برنامج فيس بوك للتحقّق من الوقائع بواسطة جهات خارجية (Lee, 2019a). وكان من بين هؤلاء الشركاء الموقع الإلكتروني Snopes الذي يكافح الخدع ويتخذ من الولايات المتحدة مقراً له. وقد أشار أحد مؤسسي Snopes.com إلى أن التحديات الأخلاقية كانت من بين أسباب الانسحاب (Green, 2019). وانسحب شريك فيس بوك في حقل التحقُّق من الوقائع في هولندا Nu.nl من المشروع أيضاً[364]. واتخذ هذا الموقع الذي لا يبغي الربح قراراً قائماً على القيم بإنهاء التعاون، اعتراضاً منه على الموقف الأخلاقي الذي تبنّاه فيس بوك والذي يعفي الإعلانات السياسية (مع بعض الاستثناءات) من التحقُّق من الوقائع (Hern, 2019a; Van Den Berg & Snelderwaard, 2019).

هناك مثال آخر على منظمة إخبارية أظهرت قواعد متنافسة، وهو الشكوى التي قدّمتها هيئة الإذاعة البريطانية BBC بشأن إعلان فيس بوك الذي استخدم مقطعاً مضللاً أُخرج من سياقه للمُحررة السياسية «لورا كوينسبرغ» تظهر فيه وكأنها تؤيد استراتيجية حزب المحافظين لخروج بريطانيا من الاتحاد الأوروبي. وقد طرح هذا الأمر تحديات هامة في مجال السمعة والأخلاقيات لهيئة إذاعة عامة تعتبر أنّ الحفاظ على الحياد السياسي يُشكّل إحدى قيمها الأساسية. وحظر فيس بوك الإعلان بعد عدة أيام

363 انظر: التحليل المفصل في الفصل 4.

364 أُعيد إطلاق برنامج فيس بوك للتحقّق من الوقائع بواسطة جهات خارجية في البلاد مع شريكَين هُما: AFP وDPA.

مـن تلقـي الشـكوى، مبـرراً ذلـك بأنـه خـرق لحقـوق الطبـع والنشـر (Mays, 2019). وعنـد حظـره، كان قـد أُنفـق مبلـغ 5000 جنيـه اسـترليني علـى الحملـة الإعلانيـة التـي ظهـرت فـي التغطيـة الإخباريـة للأحـداث الجاريـة news feed حوالـي 250,000 مـرة (Who Targets Me, 2019).

انتقـل فيـس بـوك، فـي الآونـة الأخيـرة، إلـى إبطـال الطلبـات التـي يرفعهـا ناشـرون منتسـبون سياسـياً يدّعـون أنهـم مواقـع إخباريـة محليـة ليتـم إعفاؤهـم مـن الإذن بنشـر إعـلان سياسـي الـذي فرضتـه الشـركة (Fisher, 2020). جـاء ذلـك فـي أعقـاب أبحـاث نشـرها «مركـز تـو للصحافـة الرقميـة» Tow Center for Digital Journalism وكشـف فيهـا عـن أكثـر مـن 1200 حالـة لمجموعـات سياسـية تتظاهـر بأنهـا مواقـع إخباريـة محليـة لنشـر الدعايـة فـي الولايـات المتحـدة (Bengani, 2020).

أطلقـت بعـض الـوكالات الإخباريـة التـي أدرجـت مُكافحـة المعلومـات المُضلّلـة فـي صميـم اسـتراتيجياتها التحريريـة اسـتجابات أخلاقيـة ومعياريـة إضافيـة للتضليـل. فعلـى سـبيل المثـال، حـددت دراسـة أُجريـت فـي عـام 2019 (Posetti et al., 2019a) مقاربـة «مدفوعـة بالمهمـة» لمُكافحـة المعلومـات المُضلّلـة أطلقتهـا ثـلاث منظمـات إخباريـة مـن بلـدان الجنـوب هـي: رابلـر Rappler فـي الفلبيـن، ودايلـي مافريـك Daily Maverick فـي جنـوب أفريقيـا، وكوينـت The Quint فـي الهنـد. وقـد ذكـرت كل منظمـة مـن هـذه الـوكالات الإخباريـة التزامـاً بتعزيـز المبـادئ الديمقراطيـة، والدفـاع عـن حُريّـة الإعـلام، والتمسـك بالمبـدأ الأخلاقـي الأساسـي وهـو «قـول الحقيقـة فـي مواجهـة السـلطة» رداً علـى شـبكات التّضليـل التـي ترعاهـا الدولـة ووكلاء النفـوذ الأجنبـي الذيـن اعتبـرت أنهـم يُزعزعـون اسـتقرار ديمقراطياتهـا. بالإضافـة إلـى ذلـك، سـعت إلـى تحويـل هـذه القواعـد إلـى نمـاذج تحتـذي بهـا جماهيرهـا كوسـيلة لتحفيـز المسـؤولية الأخلاقيـة عـن تفـادي ممارسـات التّضليـل، بمـا فـي ذلـك الاعتـداءات علـى الصحفييـن الذيـن يُتّهمـون «بالكـذب». وكمثـال علـى هـذه المقاربـة، يمكـن ذكـر الحمـلات المركّـزة علـى الجمهـور التـي أطلقهـا «رابلـر» والتـي تشـجّع المجتمـع علـى الانضمـام إليهـا اعتراضـاً علـى الكراهيـة عبـر الإنترنـت المرتبطـة بحمـلات تضليـل مدبّـرة اسـتهدفت «رابلـر» ورئيسـة تحريـره- مديرتـه التنفيذيـة «ماريـا ريسـا» (Posetti et al., 2019b). وقـد جـرى تفعيـل هـذه الحمـلات عبـر الإنترنـت باسـتخدام علامـات هاشـتاغ مثـل #NoPlaceForHate, #IStandWithRappler, #HoldTheLine، وكان الهـدف منهـا إظهـار التـزام أخلاقـي مشـترك بمُكافحـة المعلومـات المُضلّلـة داخـل المجتمعـات علـى الإنترنـت والوقـوف فـي وجـه حمـلات التّضليـل المسـتندة إلـى الدولـة، باعتبارهـا منافيـة للقواعـد والأعـراف الثقافيـة والاجتماعيـة.

وفـي عـام 2020، قامـت «ريسـا»، فـي معـرض النـداء الـذي وجّهتـه إلـى فيـس بـوك لتحمّـل المسـؤولية الأخلاقيـة والتصـرف اسـتجابةً للتضليـل، باقتبـاس اسـتنتاج توصّلـت إليـه الأمـم المتحـدة ومفـاده أن فيـس بـوك أدى «دوراً حاسـماً[365]» فـي مـا وصفتـه الأمـم المتحـدة بأنـه «مثـال نموذجـي عـن التطهيـر العِرقـي[366]» ضـد «الروهينغـا» فـي ميانمـار مـن خـلال عـدم الاكتـراث بالمعلومـات المُضلّلـة وخطـاب الكراهيـة (Posetti, 2020). واعتـرف فيـس بـوك فـي وقـت لاحـق بمـا يلـي «لـم نكـن نفعـل مـا يكفـي للمسـاعدة فـي منـع اسـتخدام منصتنـا لـزرع الانقسـام والتحريـض علـى العنـف خـارج الإنترنـت»، مضيفـاً أنـه قـام بتحديـث سياسـاته «لإزالـة المعلومـات المغلوطـة القـادرة علـى المسـاهمة فـي العنـف الوشـيك أو الأذى الجسـدي[367]».

365 https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_CRP.2.pdf

366 https://news.un.org/en/story/2017/09/564622-un-human-rights-chief-points-textbook-example-ethnic-cleansing-myanmar

367 https://about.fb.com/news/2018/11/myanmar-hria/

من جهته، اضطلع «المنتدى الوطني للمحررين» في جنوب أفريقيا أيضاً بدور معياري من خلال تسليط الضوء على مخاطر التّضليل والعمل بشكل وثيق مع هيئة رصد الإعلام في أفريقيا، تأكيداً على الالتزام باستئصال المعلومات المُضلّلة[368]. وتشمل المبادرات المتصلة بهذا التعاون بوابة الإبلاغ عن المعلومات المُضلّلة Real411.

برزت استجابة تعاونية أخرى للتضليل من مجتمع الصحافة خلال القمة العالمية للإعلام في عام 2019 (وهي مبادرة أطلقتها وكالة أنباء Xinhua الصينية، وتضم حالياً 13 شريكاً إعلامياً دولياً من بينهم رويترز، وهيئة الإذاعة البريطانية BBC، وAP). وتفيد التقارير بأن القمة توصّلت إلى توافق آراء بشأن المعلومات المُضلّلة: «لضمان احترام سلطة الإعلام ومصداقيته، يتعيّن على وسائل الإعلام مُكافحة المعلومات المُضلّلة؛ وينبغي توضيح المعلومات المغلوطة دون تأخير؛ ولابدّ من مقاطعة الجميع للأخبار الزائفة... إنّ الإبلاغ عن أخبار زائفة ونشرها ينتهكان الأخلاقيات الصحفية ويضرّان بمصالح الجمهور» (Xinhua, 2019).

د. المبادرات المناهضة لخطاب الكراهية:

في الحالات التي تتقاطع فيها المعلومات المُضلّلة مع خطاب الكراهية - مثل العنصرية وكراهية النساء والتعصّب، كثيراً ما تنشأ استجابات معيارية وأخلاقية تتراوح بين مبادرات تطلقها منظمات المجتمع المدني والوكالات الحكومية الدولية وتدخلات من المشاهير. وأحد هؤلاء المشاهير هو الممثل الكوميدي «ساشا بارون كوهين» Baron Cohen Sacha، الذي سعى في الخطاب الذي ألقاه حول التّضليل والدعاية اللذين تغذيهما وسائل التواصل الاجتماعي خلال المؤتمر الذي نظّمته رابطة مناهضة التشهير حول معاداة السامية والكراهية في تشرين الثاني/نوفمبر 2019 إلى حثّ الشركات المشاركة على اتخاذ إجراءات ضد المعلومات المُضلّلة التي تعرّض الأقليات الدينية والإثنية للخطر (Baron Cohen, 2019).

تُسجَّل أيضاً تدخّلات من جانب معاهد أبحاث ومنظمات غير حكومية سعياً منها إلى توفير التوجيه المعياري من خلال وضع أُطُر مصمّمة لترسيخ مقاربات قائمة على القيم لإدارة خطاب الكراهية، الذي يبرز كسِمة من سمات حملات التّضليل المدبّرة. ومن الأمثلة على هكذا تدخلات؛ قيام «المؤسسة الدولية للنظم الانتخابية» باستكشاف الروابط القائمة بين خطاب الكراهية والتّضليل، وتوفير إطار معياري لتدخلات برمجية (Reppell & Shein, 2019).

ويتوفّر مثال آخر في هذا السياق، وهو تقرير لمنظمة «مُراسلون بلا حدود» بعنوان «التنمُّر الإلكتروني ضد الصحفيين» Attack of the Trolls الذي يغطّي الهجمات التي يتعرّض لها الصحفيون عبر الإنترنت -لا سيما عند تقاطع كراهية النساء والتّضليل. وقد سعى التقرير إلى نشر الوعي وتفعيل الاستجابات المصمّمة لتعزيز قواعد حُريّة الصحافة عبر الإنترنت (RSF, 2018) وبالمثل، ركّزت مبادرة #JournoDefender initiative[369] التي أطلقها معهد FOJO للإعلام على مُكافحة كراهية النساء عبر الإنترنت لأنها تتقاطع مع هجمات التّضليل المصمّمة لتقويض الديمقراطية. وقد استندت هذه المبادرة إلى أبحاث أُجريت في بلدان متعددة (FOJO: Media Institute, 2018).

....

[368] https://sanef.org.za/disinformation/

[369] https://journodefender.org/

سعى الصحفيون والوكالات الإخبارية إلى تعزيز قيم المساواة بين الجنسَين من خلال التحقيق في حملات التّضليل التي تشمل عناصر كراهية النساء. فعلى سبيل المثال، ذكرت «ماريا ريسا»، رئيسة تحرير Rappler ومديرته التنفيذية الالتزام بمبدأي «قول الحقيقة في مواجهة السلطة» و«تسليط الضوء» كأسباب دفعتها إلى التحدث علناً عن تجربتها لفائدة التعرض لمضايقات وحشية على الإنترنت انتقاماً من الصحافة الاستقصائية التي كشفت عن شبكات تضليل ترتبط بالحكومة في الفلبين (GIJN Staff, 2019).

أخيراً، أشار المقررون الخاصون للأمم المتحدة إلى قيام خطاب الكراهية عبر الإنترنت باستخدام أساليب التّضليل ضد صحفيات. وقد أصدر خمسة مقررين خاصين للأمم المتحدة بياناً مشتركاً في عام 2018 يدعو الحكومة الهندية إلى حماية الصحفية الهندية «رنا أيوب» التي وُجّهت إليها تهديدات بالقتل في إطار حملة تضليل تطبعها كراهية النساء استخدمت مقاطع فيديو «زائفة بفجاجة» وحسابات زائفة لوصفها بصورة مضللة وتعريضها للخطر (UN Human Rights, 2018).

4.1.7 دراسة حالة بشأن الاستجابة: معلومات مضللة حول جائحة فيروس كورونا

تشمل الاستجابات الأخلاقية والمعيارية الإدانة العلنية لأعمال التّضليل أو التوصيات والقرارات الرامية إلى إحباط هذه الأفعال بسبب طبيعة الوباء المهدّدة للحياة. وتشمل هذه الاستجابات بيانات من المقررين الخاصين للأمم المتحدة ومسؤولي منظمة الصحة العالمية والقادة الوطنيين. بالإضافة إلى ذلك، توفّرت أمثلة على دعوات لتعزيز السلوك الأخلاقي داخل عالم الصحافة، وطلب وُجّه إلى شركات ومنصات التواصل عبر الإنترنت لاعتماد تدابير إضافية. وقد اتخذت هذه الاستجابات، في كثير من الأحيان، شكل بيانات أو خطابات أو مقالات منشورة ترمي إلى دفع الآخرين إلى التوقف عن مشاركة معلومات مضللة، وتعزيز قواعد حُريّة التّعبير خلال الأزمة، وتكييف المعايير الأخلاقية لمواجهة التحديات الجديدة في الاستجابة لما أسماه موجزان للسياسات؛ طلبت اليونسكو إجراءهما استجابةً لأزمة المعلومات المُضلّلة حول فيروس كورونا «مُشكلة التّضليل» (Posetti & Bontcheva, 2020a; Posetti & Bontcheva, 2020b).

وتشمل الأمثلة ما يلي:

- بيان أصدره في اليوم العالمي لحُريّة الصحافة الأمين العام للأمم المتحدة «أنطونيو غوتيريش» يعزّز الدور المعياري للصحافة المهنية في النظام البيئي للمعلومات كحصن ضدّ المعلومات المُضلّلة. وقد أكّد هذا البيان أيضاً على الالتزامات الأخلاقية والقانونية للدول الأعضاء في الأمم المتحدة في ما يتعلق بحقوق حُريّة الصحافة (وآليات سلامة الصحافة) في سياق الاستجابات لجائحة فيروس كورونا (UN Secretary General, 2020).
- بيان مشترك صادر عن خبراء دوليين، من بينهم «دايفيد كاي»، مقرر الأمم المتحدة الخاص المعني بالحق في حُريّة الرأي والتّعبير؛ و«هارلم ديزير»، ممثل منظمة الأمن والتعاون في أوروبا المعني بحُريّة الإعلام؛ و«إديسون لانزا»، المقرر الخاص للجنة البلدان الأمريكية لحقوق الإنسان المعني بحُريّة التّعبير: «يجب على الحكومات تعزيز الوصول إلى المعلومات والتدفق الحر للمعلومات أثناء الوباء وحمايتها» (UN Human Rights, 2020a).

- تقرير رفعه إلى مجلس حقوق الإنسان التابع للأمم المتحدة المقرر الخاص للأمم المتحدة المعني بالحق في حُريّة الرأي والتّعبير «دايفيد كاي»، ناشد فيه صراحةً الدول الأعضاء التمسّك بالالتزامات والمسؤوليات الأخلاقية في ما يتعلّق باستجاباتها لفيروس كورونا (Kaye, 2020a). ويذكر التقرير أنه «نداء إلى جميع الحكومات لمعاملة مَن هم في نطاق ولاياتها القضائية... بالكرامة والاحترام اللذين يتطلبهما القانون الدولي لحقوق الإنسان».
- دعوات وجّهها كبار المحررين والصحفيين والأكاديميين الإعلاميين لوقف البث المباشر لخطابات السياسيين الذين ينشرون معلومات مضللة ولمؤتمراتهم الصحفية، وذلك بسبب صعوبة التحقُّق من الوقائع وكشف زيفها في الوقت الحقيقي (Thomas, 2020).
- قرارات غير مسبوقة صادرة عن شركات التواصل عبر الإنترنت لتعديل تسجيلات للقادة السياسيين الذين يُعتبر أنهم ينشرون معلومات مُضلّلة حول جائحة فيروس كورونا أو لإزالتها (BBC, 2020b).
- كما أُشير إليه أعلاه، أدت الأزمة إلى قيام أكثر من 130 دولة عضو في الأمم المتحدة ومراقبين رسميين بالمطالبة بإسناد جميع الخطوات الرامية إلى التصدي لجائحة فيروس كورونا، من بين جملة أمور، إلى احترام حُريّة التّعبير وحُريّة الصحافة[370].
- أصدرت اللجنة الأفريقية لحقوق الإنسان والشعوب بياناً صحفياً بشأن «الاستجابة الفعّالة القائمة على حقوق الإنسان» للوباء. وكرّر هذا البيان مرةً أخرى واجب الدول في ضمان امتثال التدابير المعتمدة لمبدأ الشرعية، وأهمية أن تكون ضرورية ومتناسبة مع هدف حماية الصحة العامة. وتشمل هذه التدابير تلك التي تهدف إلى «التخلّص من المعلومات المغلوطة والخرافات حول جائحة فيروس كورونا والمعاقبة على نشر معلومات كاذبة عن مخاطر الفيروس[371]».

5.1.7 كيف يجري تقييم الاستجابات الأخلاقية والمعيارية؟

هناك أدلة محدودة للغاية على تقييم الاستجابات الأخلاقية والمعيارية للتضليل، ويعود ذلك جزئياً إلى الصعوبة المنهجية التي تنطوي عليها هذه العملية. ومن العوامل المساهمة في ذلك أن ترسيخ الأخلاقيات والقواعد داخل المجتمعات، أو تحفيز الالتزامات بالمبادئ الدولية لحقوق الإنسان لعملية تعاونية إلى حد كبير، ومن شبه المستحيل تحديد جهة فاعلة أو رسالة معينة كانت تحويلية إلى حد ما.

يرصد المقرر الخاص للأمم المتحدة ذو الصلة تعاطي الدول الأعضاء مع التّضليل، في إشارة إلى تمسّكها بالقواعد الدولية لحقوق الإنسان مثل حُريّة التّعبير، ويرفع تقييمات إلى مجلس حقوق الإنسان التابع للأمم المتحدة على هذا الأساس (UN Human Rights, 2020b). ولكن لا توجد عملية تقييم معروفة تسعى مباشرةً إلى إسناد تطوير القواعد والأخلاقيات داخل المجتمعات إلى مثل هذه التدخلات. غير أن بعض مراجع الحالات تظهر تأثيرها في سياقات معينة[372].

370 https://www.un.org/africarenewal/news/coronavirus/cross-regional-statement-%E2%80%9Cinfodemic%E2%80%9D- con-text-covid-19

371 https://www.achpr.org/pressrelease/detail?id=483

372 https://www.lefigaro.fr/flash-actu/la-bolivie-abroge-des-decrets-anti-desinformation-controverses-20200515

فعلى سبيل المثال، تقوم وسائل الإعلام الإخبارية إجمالاً بتضخيم أي بيان يصدر عن الأمين العام للأمم المتحدة (UN Secretary General, 2020) ويسلّط فيه الضوء على ضرورة ألا تؤدي تدابير مُكافحة التّضليل التي تتخذها الدول الأعضاء إلى تقويض مبادئ مثل حُريّة الصحافة (Apelblat, 2020) فيما تقوم منظمات المجتمع المدني بتعزيزه لترسيخ قواعد مثل «النفاذ إلى المعلومات» (Article 19, 2020). ولكن، إذا ما ذهبنا أبعد من عمليات القياس الإعلامي التي يجريها مقدمو الخدمات التجارية، لا نجد أي أبحاث منهجية واسعة النطاق متاحة للجمهور تتمحور حول مدى نشر هذا النوع من البيانات وتضخيمها.

6.1.7 التحديات والفرص

يمكن أن يكون تطبيق هذه التدخلات المعيارية والأخلاقية بسيطاً نسبياً وغير مكلف، ويمكن أن تعمل هذه التدخلات بمثابة سرديات مضادة تجذب البوصلات الأخلاقية للأفراد، أو تعزز احترام قيم مثل مناهضة العنصرية أو مناهضة كراهية النساء. بيد أن هناك مشكلة تتمثل في عدم ارتباط البوصلات الأخلاقية والقواعد المجتمعية بمبادئ الوصول إلى المعلومات وحُريّة التّعبير وحُريّة الصحافة والخصوصية –على النحو المنصوص عليه في القانون الدولي لحقوق الإنسان. وتبرز محاولات عديدة لتطبيع التّعبير عن المُحتوى الكاذب والمضلل الذي قد يكون ضاراً ونشره.

تشكّل المقاربات القانونية والتنظيمية التي تتعارض والقواعد الدولية لحُريّة التّعبير (بما في ذلك حُريّة الصحافة ذات الصلة) والخصوصية أحد أهم المخاطر المرتبطة باستجابات الدول للتضليل.

وكما أظهر الاندفاع نحو الاستجابة لمشكلة التّضليل التي رافقت فيروس كورونا، تستطيع أزمة المعلومات المضللة أن تقود إلى تغيير ما يُعتبر طبيعياً ومقبولاً، مثل تعليق حماية حقوق الإنسان أو إضعافها (Posetti &Bontcheva 2020a; Posetti & Bontcheva 2020b). علاوةً على ذلك، وعلى الرغم من أن القيود المفروضة على الحق في التماس المُحتوى وتلقّيه ونقله قد تكون مشروعة بموجب المعايير الدولية، يبقى أنه متى فُرضت هذه القيود لأسباب تتعلق بالصحة العامة، ينبغي أن ينصّ عليها القانون وأن تكون ضرورية ومتناسبة مع الغرض منها.

في مثل هذه الظروف، يمكن أن تتجاوز الإستجابات – كالقوانين الجديدة الخاصة «بالأخبار الزائفة» والتي تجرّم الصحافة بشكل فعّال– هذه المعايير، وتترسّخ بالتالي كقواعد جديدة. لهذا السبب، يشكّل الحرص على أن تكون جميع التدخلات مرتكزة إلى الأطر القانونية والمعيارية لحقوق الإنسان، لا سيما حُريّة التّعبير (بما في ذلك حُريّة الصحافة والوصول إلى المعلومات) واحترام الخصوصية، تحدياً أكيداً.

في نهاية المطاف، قد يقوم الأفراد والمنظمات والدول المعارضة للنية وراء الاستجابات المعيارية والأخلاقية المشروعة للتضليل بنزع صفة الشرعية عنها، تماماً كما يقوم أولئك الذين يسعون إلى تجنّب المحاسبة بوصف الصحافة ذات المصداقية بشكل مضلل على أنها «أخبار زائفة».

تتصدّى جهات فاعلة عديدة تسلّط الضوء على هذه القضايا للتحدي المتمثل في خطر خفض رتبة معايير حقوق الإنسان من خلال تمكين الجمهور (وممثليهم المنتخبين) من الاعتراف بضرورة تحديد مثل هذه التدخلات ضد التّضليل زمنياً (حيث تكون مبررة أثناء حالات الطوارئ، مثل جائحة فيروس

كورونا)، من بين جملة أمور. ومع ذلك، يعتمد تأثير هذه الرسائل على إقناع الجهات التي تتمتع بالسلطة بالالتزام عن كثب بهذه المعايير.

وتتمثّل الفرصة الرئيسية في الاستجابات الأخلاقية والمعيارية للتضليل في إعادة التأكيد على القواعد المتعلقة بالوصول إلى المعلومات وحُريّة التّعبير وبتذكير الناس بها. وفي خضمّ جائحة فيروس كورونا، يمكن التأكيد على أن هذه القواعد لا تطال فحسب الحقوق الأساسية، بل تشكّل أيضاً أدوات مهمّة للتخفيف من تأثير المعلومات المُضلّلة والتصدي لها.

ويمكن أيضاً اتخاذ خطوات معيارية فورية للتصدي للتضليل بغرض تعزيز الآثار المعيارية والمؤسسية طويلة الأجل للمعايير الدولية. فعلى سبيل المثال، يمكن أن تفسّر التقارير الإخبارية عن الاستجابات للتضليل أهمية تقييم هذه المبادرات على ضوء القوانين الدولية لحقوق الإنسان، والأُطُر المعيارية والأخلاقية التي تدعمها. ويمكن أن تساعد هذه الصحافة التفسيرية على تعزيز المساءلة من جانب الحكومات والشركات، كما تساعد على ترسيخ فهم الدور الذي تؤديه هذه القيم والمعايير في مجالات تتجاوز التصدي للمعلومات المُضلّلة.

7.1.7 توصيات بشأن الاستجابات المعيارية والأخلاقية

توفّر التحديات والفرص المحددة أعلاه، وتبعاتها الهامة على حُريّة التّعبير، أساساً لخيارات العمل في هذه الفئة من الاستجابات.

تشمل التوصيات المتعلقة باتخاذ إجراءات في هذا القسم تعزيز الأسس المؤسسية لقواعد حُريّة التّعبير، باعتبارها عناصر من الاستجابات للتضليل. وفي هذا الصدد، يمكن رفع التوصيات التالية:

توصيات للمنظمات الدولية:

- إجراء تقييم متابعة لتداول البيانات المعيارية والانخراط فيها، فضلاً عن تقييم الأثر الفعلي لمدونات الأخلاقيات، كالتي تطبقها شركات ومنصات التواصل عبر الإنترنت ووسائل الإعلام الإخبارية ذات الصلة بقضايا المعلومات المُضلّلة.

توصيات للدول الفردية:

- ضمان الحوكمة المؤسسية الشاملة لأصحاب مصلحة متعددين لشركات ومنصات التواصل عبر الإنترنت، بما يشمل الشفافية والسياسات المتعلقة بالتّضليل في سياق تنظيم المُحتوى.
- إدراج تقييمات التأثير على حقوق الإنسان ضمن الاستجابات للمعلومات المُضلّلة التي تعتمدها الفروع التنفيذية أو التشريعية للحكومة، خاصةً تلك التي قد يكون نطاقها مفرطاً (مثل توسيع نطاق قوانين «الأخبار الزائفة» في سياق جائحة فيروس كورونا).

توصيات لشركات ومنصات التواصل عبر الإنترنت:

- الالتزام بالقيم التي تدافع عن المجتمعات المحلية والفئات الضعيفة، بما في ذلك التهديدات بلغات متعددة، وضمان أن تتوفّر تدابير لمُكافحة التّضليل في جميع البلدان التي تعمل فيها.
- إشراك أصحاب مصلحة متنوّعين في وضع سياسات تدعم اتخاذ القرارات الأخلاقية في مجال المعلومات المُضلّلة- بما في ذلك ما إذا كان ينبغي إزالتها.
- زيادة القدرة على التعامل مع التّضليل على نطاق واسع، لا سيما في البلدان التي تشهد عدم استقرار و/أو نزاعات وقلاقل، وتوفير استجابات سريعة للجهات الفاعلة التي تستهدفها هذه المعلومات المُضلّلة، فضلاً عن فرص الإنصاف بعد صدور قرارات حول كيفية التعامل مع المُحتوى ذي الصلة.
- تعزيز دورها المعياري في مجال حُريّة التّعبير عن طريق ضمان إجراء مراجعة مستقلة ومنتظمة لسياساتها المتعلقة بالتّضليل، وما يترتب على ذلك من آثار على حقوق الإنسان، وتنفيذها.
- الاعتراف بأن الالتزام الأخلاقي بحُريّة التّعبير لا يحول دون اتخاذ مجموعة من الإجراءات الحاسمة المتعلقة بالتّضليل السياسي الذي قد يتسبّب بضرر كبير- كالحالات التي يُهدد فيها الحياة أو الصحة العامة أو مؤسسات ديمقراطية أو النزاهة الانتخابية.
- تعزيز الشفافية والكشف عن البيانات المتعلّقة بالعمليات الآلية إلى إدارة المعلومات المُضلّلة.

توصيات للجهات الفاعلة الإعلامية:

- الحرص على التزامها بأعلى المعايير الأخلاقية والمهنية لتجنب الوقوع في أيدي المروّجين للتضليل أو ربطها بهم.
- الاستثمار في الصحافة الاستقصائية التي تركّز على كشف شبكات المعلومات المُضلّلة وشرح مخاطرها لجماهيرها، وأهميّة مقاومتها من أجل المصلحة العامة، كوسيلة لبناء الثقة مع السعي إلى تأمين الحقيقة.
- زيادة قدرة مجالس الصحافة المستقلة على رصد المعلومات المُضلّلة والتصدي لها (بما في ذلك عند انتشارها عبر قنوات إعلامية إخبارية) والاستجابات للتضليل (خاصة لأنها تؤثر على حُريّة التّعبير) في إطار دورها الرقابي الأخلاقي.

توصيات للباحثين:

- استخدام أساليب البحث لقياس تأثير الرسائل الرامية إلى تطوير الأخلاقيات والقيم التي تُساعد على التحصين ضد المعلومات المُضلّلة، أو إجراء أبحاث نوعية بشأن التطور المعياري وتغيّر السلوك استناداً إلى التصدي للتضليل.
- دراسة مبادرات الدراية الإعلامية والمعلوماتية لتقييم التأثير على سلوكيات المشاركين وشعورهم بالمُساءلة الشخصية لفائدة التصدي للتضليل.

2.7 الاستجابات التثقيفية

المؤلفون: كالينا بونتشيفا، وجولي بوزيتي، ودينيس تيسو

تهدف الاستجابات التثقيفية إلى تحسين دراية المواطنين في مجال الإعلام والمعلومات، وتعزيز التفكير الناقد والتدقيق عند الحصول على المعلومات عبر الإنترنت، فضلاً عن التدريب الصحفي واستخدام الأدوات المصممة لتعزيز التحقُّق من الوقائع والتدقيق وكشف زيف المعلومات.

تكتسب أهميةً خاصةً في هذا الصدد مسائل التفكير الناقد والدراية الإخبارية والإعلانية، والتوعية بحقوق الإنسان، وقضايا الهُويّة، وفهم الخوارزميات والبيانات، ومعرفة الاقتصاد السياسي للتواصل (بما في ذلك الاقتصاد وعلم النفس وعلم الاجتماع التي تحيط كلّها بإنتاج معلومات مُضلّلة واستهدافها ونشرها).

يقدم هذا القسم لمحة شاملة عن أنواع الاستجابات التثقيفية المختلفة، ويقوم بذكر المنظمات التي تصممها وتنفذها، ويقوم كذلك بتحديد أهداف هذه الاستجابات. وفي الختام، يتم تقييم الاستجابات هذه استناداً إلى كيفية تصديها للتضليل من خلال تثقيف المتعلمين بشأن القيمة الأساسية لحُريّة التّعبير، وشرح الفرق بين تعبئة وقائع مختلفة وتفسيرها من جهة (وهي لا تشكّل التّضليل)، وتعبئة معلومات مغلوطة أو مضللة وتفسيرها من جهة أخرى (علماً أن هذا الأمر هو جوهر التّضليل).

1.2.7 ما ومن الذي تستهدفه الاستجابات التثقيفية؟

تُعتبر مبادرات الدراية الإعلامية والمعلوماتية والتفكير الناقد من مهارات القرن الحادي والعشرين الرئيسية، وهي مهارات يحتاج اليها المواطنون لتمييز المعلومات المُضلّلة على الإنترنت والتصدي لها بشكل أكثر فعاليةً. وكما أشار تقرير صادر عن لجنة النطاق الواسع: «تميل المناهج الدراسية التقليدية إلى إعطاء الأولوية لتراكم المعرفة على حساب تطبيقها، وتفشل نظم مدرسية كثيرة في تدريب الطلاب بشكل مناسب على المواطنة والإلمام الرقميين» (Broadband Commission, 2013).

يشمل مفهوم الدراية الإعلامية والمعلوماتية، كما تستخدمه اليونسكو، مجموعة من الكفايات المتعلقة باستخدام المُحتوى وإنتاجه وتداوله. وترد تحت هذه المظلة المعرفة والمهارات التي تغطي مجالات مثل التفكير الناقد؛ وإنتاج المُحتوى؛ والدراية الإخبارية؛ والدراية الإعلانية؛ والدراية السينمائية؛ والاقتصاد السياسي للتواصل؛ والدراية الخوارزمية؛ والدراية بالخصوصية؛ والتواصل بين الثقافات (Berger, 2019). وتستخدم اليونسكو أيضاً مفهوم التعليم من أجل المواطنة العالمية[373] الذي يشمل الكفايات المحيطة بالهُويّة والقيم. وتمثل هذه المفاهيم معاً قواعد اللُّعبة التي تساعد في تمكين المشاركين في مجال التواصل الرقمي من التعامل مع المعلومات المُضلّلة بمجموعة من الطرق. وقد تكون المبادرات التثقيفية في المجال الواسع للدراية الإعلامية والمعلوماتية نظامية وغير نظامية، وهي تنتشر عبر مجموعة من المؤسسات الاجتماعية؛ انطلاقاً من المدارس ووصولاً إلى المدن، وعبر أنظمة النقل، علاوةً على وسائل الإعلام ووسائل التواصل الاجتماعي.

373 https://en.unesco.org/themes/gced

تشـدّد مجموعـة اليونسـكو الواسـعة مـن الكفايـات المسـتهدفة علـى النطـاق الشـامل الـلازم لنجـاح الجهـود فـي مجـال الدرايـة الإعلاميـة والمعلوماتيـة. فـي حيـن تميـل جهـود كثيـرة إلـى التركيـز علـى الدرايـة بالأخبـار والتدقيـق فقـط، نـادراً مـا تُعالـج مواضيـع الدرايـة الخوارزميـة والإعلانيـة والدرايـة بالخصوصيـة شـديدة الترابـط. ويُصـاغ مفهـوم «الإلمـام الرقمـي» بشـكل مختلـف وفقـاً للكفايـات التـي يهـدف إلـى وصفهـا. وقـد اعتبـرت بعـض منظمـات المجتمـع المدنـي[374] أنـه مـن الأهميـة بمـكان أيضـاً تثقيـف الأطفـال (والكبـار) بشـأن أسـلوب جمـع البيانـات الشـخصية ومشـاركتها عبـر الإنترنـت لتحقيـق مكاسـب تجاريـة؛ والأخطـار الخفيـة لتحديـد السـمات والاسـتهداف عبـر الإنترنـت؛ والخوارزميـات وتحيزاتهـا؛ وخصوصيـة المسـتخدم علـى الإنترنـت. وتبـرز بالتالـي ثغـرة هامـة فـي عـدد مـن مجموعـات الأدوات والبرامـج المتعلقـة بالدرايـة الإعلاميـة والمعلوماتيـة تتمثـل فـي عـدم تغطيـة مفهـوم الإلمـام بالبيانـات. وفـي الوقـت نفسـه، يرتبـط الإلمـام بالبيانـات فـي مواجهـة الروابـط الـى المعلومـات المُضلّلـة ارتباطـاً وثيقـاً بالمواطنـة الرقميـة (Carmi et al., 2020).

تشكّل معالجة هذه التحديات المرتبطة بالدراية الإعلامية والمعلوماتية من خلال استجابات تثقيفية فعالة وطويلة الأمد جزءً رئيسياً من الأحجية، ذلك أن الأبحاث وجدت في بعض الحالات أن العامل الرئيس على تضخيم المعلومات المُضلّلة واسعة الانتشار هو العنصر البشري (Vosoughi et al., 2018). وتُطرح أسئلة رئيسية في هذا الصدد لمعرفة لماذا «يقع المواطنون» في فخ التّضليل عبر الإنترنت، وما الذي يحفّزهم على مشاركة معلومات مضللة (حتى وإن كانوا يدركون أنها غير صحيحة)، وما هو تأثير المعلومات المُضلّلة عبر الإنترنت على سلوكهم خارج إطار الإنترنت (هل يؤثر ذلك مثلاً على تصويتهم في الانتخابات)؟ وفي سياق جائحة فيروس كورونا على وجه التحديد، يتعرّض الكثير من المواطنين للخداع فينشرون معلومات مضللة عبر الإنترنت، مما يجعلهم غير قادرين على فهم التدابير الوقائية القائمة على أُسُس علمية وعلى تنفيذها. ويلقى الناس حتفهم نتيجةً للتراخي (Karimi & Gambrell, 2020) أو اللُّجوء إلى «وصفات» كاذبة (Embury-Dennis, 2020).

انكبّ كلّ من العلماء (مثل Corbu et al., 2020) والجهات المعنية بالتحقُّق من الوقائع (Vicol, 2020) على دراسة السبب الذي يدفع المواطنين إلى تصديق مُحتوى مغلوط أو مضلل وإلى نشره. وقد تم التأكيد على أن العمر والتعليم والتعرّض المتكرر للتضليل هي كلّها عوامل مهمة (Vicol, 2020)، علماً أن البالغين الذين تتجاوز أعمارهم الخمسين والذين يفتقرون إلى مستوى تعليمي عالٍ معرّضين للخطر بشكل خاص. ويتمثّل عامل مهم آخر في الانحياز التأكيدي، أي ميل الناس لقراءة المُحتوى الذي يتوافق مع وجهات نظرهم العالمية القائمة، وتصديقه (Nickerson, 1998;Corbu et al.,2020; Nygren & Guath, 2019). وفقاً لدراسة أجراها «غالوب» و«مؤسسة نايت» Knight (Knight Foundation, 2018)، يتبادل الناس عموماً المعلومات التي يثقون بها ويفعلون ذلك في المقام الأول لأسباب اجتماعية أو شخصية. علاوةً على ذلك، يتأثر سلوك الفرد عبر الإنترنت في مجال مشاركة الأخبار والمعلومات والتعليق عليها بسلوك روابطه الاجتماعية المتشابهة في التفكير عبر الإنترنت– ويُشار إلى هذه الفئة في الدراسات العلمية بعبارة homophily أو فئة المتماثلين (Tarbush & Teytelboym, 2012). ومما لا شكّ فيه أن تلقّي المُحتوى من مصادر موثوقة، كالأصدقاء والعائلات، يعزّز مصداقية هذا المُحتوى.

وجـد باحثـون أن «الشـبكات الاجتماعيـة ومحـركات البحـث ترتبـط بارتفـاع فـي المسـافة الأيديولوجيـة المتوسـطة بيـن الأفـراد» (Flaxman et al., 2018)، أي تـؤدي إلـى الاسـتقطاب، وقـد تكـون هـذه المسـألة ذات صلـة أيضـاً. وقـد بـرزت هـذه النتائـج فـي الكثيـر مـن البلـدان (Kelly & Francois, 2018).

374 https://5rightsfoundation.com/our-work/data-literacy/

وأظهرت الأبحاث التجريبية أيضاً أنه عندما تتعرّض المجتمعات المستقطبة عبر الإنترنت لمعلومات مضللة تتوافق مع السرديات المفضلة لديها، تصدّقها وتشاركها (Quattrociocchi et al., 2016). وبالتالي، عندما يواجه هؤلاء المستخدمون وهذه المجتمعات كشف زيف المعلومات أو آراء متعارضة، قد يترك هذا الأمر تأثيراً ضئيلاً عليهم أو قد يُعزّز حتّى معتقداتهم ومفاهيمهم الخاطئة القائمة من قبل. علاوةً على ذلك، أظهر مسح أجرته منظمة FullFact مؤخراً أنّ فئة المتماثلين حفّزت 25% من البالغين في المملكة المتحدة على مشاركة مُحتوى، على الرغم من أنهم كانوا يعتقدون أنه مُختلق أو مبالغ فيه (Vicol, 2020).

ووجد باحثون من مشروع YouCheck! بشأن الدراية الإعلامية والمعلوماتية أيضاً أدلة على الإفراط في الثقة (Nygren & Guath, 2019; Nygren & Guath, 2020) وعلى «نظرية تأثّر الآخرين بالرسائل الإعلامية» (Durach, 2020; Corbu et al., 2020)، حيث أن الناس يقيّمون قدراتهم الخاصة على الكشف عن معلومات مضللة على الإنترنت بشكل إيجابي للغاية مقارنةً بقدرات الآخرين.

وتشير هذه الأدلة مجتمعةً إلى ضرورة استحداث استجابات فعّالة تستند إلى الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية للتصدي للتضليل.

2.2.7 من الذي تُحاول الاستجابات التثقيفية مساعدته؟

تُعتبر الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية من المهارات الرئيسية التي تمكّن المواطنين من تمييز المعلومات المُضلّلة عبر الإنترنت بشكل أكثر فعالية. غير أنه في الحالات التي أُجريت فيها دراسات استقصائية موجّهة نحو المواطنين، بيّنت أدلة ثابتة أن غالبية المواطنين يفتقرون إلى هذه المهارات الأساسية. فعلى سبيل المثال، أثبت مسح Eurobarometer لعام 2018 (Eurobarometer, 2018) الذي أُجري في الدول الأعضاء في الاتحاد الأوروبي البالغ عددها 28 دولة أن 15٪ فقط من المجيبين شعروا بثقة شديدة في تحديد المعلومات المُضلّلة عبر الإنترنت. وركّزت دراسات استقصائية أخرى بشكل خاص على قدرة المواطن على التمييز بين البيانات الوقائعية وبيانات الرأي، وكشفت دراسة لـ«مركز بيو Pew للأبحاث» (Mitchell et al., 2018) أن 26٪ فقط من الأمريكيين (بشكل متوسط) تمكنوا من التعرف إلى البيانات الإخبارية الوقائعية، مع ارتفاع هذه النسبة لتبلغ 33٪ في صفوف الأمريكيين الأصغر سناً. إلى جانب ذلك، أظهرت نتائج تقرير RISJ (Newman, 2019b) ميلاً عالمياً إلى الخلط بين الصحافة السيئة والتّضليل و«الأخبار الزائفة».

وقد أدت هذه المسألة إلى ظهور مبادرات تسعى إلى تحسين الدراية الإعلامية والإلمام الرقمي والإلمام بالبيانات والتفكير الناقد في جميع الأعمار (من تلاميذ المدارس إلى المتقاعدين). وترتبط مسألة الإلمام بالبيانات في مواجهة المعلومات المُضلّلة ارتباطاً وثيقاً بالمواطنة الرقمية (Carmi et al, 2020).

تُستكمل هذه الجهود بمبادرات وموارد تهدف إلى تثقيف طلاب الصحافة والصحفيين المحترفين بأحدث الأدوات والمنهجيات والموارد للتدقيق في المعلومات المُضلّلة عبر الانترنت والتحقيق فيها وكشف زيفها. وغالباً ما يتم تطويرها وتيسيرها من قبل كبار الصحفيين، والمُدربين الصحفيين، والباحثين، ومنظمات المجتمع المدني. وكثيراً ما تكون هذه الجهود تعاونية إلى حد كبير.

3.2.7 ما هي المخرجات التي تنشرها الاستجابات التثقيفية؟

مُخرجات تهدف إلى تحسين الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية لدى المواطن:

تعتمد إحدى فئات مبادرات الدراية الإعلامية على التعلم من خلال **الألعاب عبر الإنترنت**، أي تعليم المواطنين الدراية الإعلامية والتفكير الناقد من خلال المشاركة في لعبة. إنها لطريقة تحفّز الناس (وليس فقط أطفال المدارس) على اكتساب المعرفة والخبرة. ومن الأمثلة على ذلك، مبادرة Drog [375] التي جمعت أكاديميين وصحفيين وخبراء إعلاميين لاستحداث لعبة على الإنترنت اسمها GetBadNews. تهدف اللعبة إلى تثقيف الناس بشأن التكتيكات المختلفة المستخدمة في حملات الدعاية والتّضليل على الإنترنت. وهناك لعبة تثقيفية أخرى اسمها Fakey [376] طوّرتها جامعة إنديانا تطلب من اللاعبين مشاركة مقالات ذات مصداقية أو وضع نقرة إعجاب عليها والإبلاغ عن تلك المشبوهة ليتم التحقُّق من وقائعها. وقد طوّرت هيئة الإذاعة البريطانية BBC لعبة تفاعلية أسمتها iReporter (Scott, 2018) [377] يؤدي في إطارها اللاعبون الشباب دور الصحفيين الذين يتعيّن عليهم الإبلاغ عن الأخبار دون الوقوع فريسة للتضليل. ويتوفّر مثال بارز آخر هو مشروع YouCheck! متعدد اللغات.

Detectives [378] هي لعبة حول الأخبار الزائفة متوفرة بالإنكليزية والفرنسية والإسبانية والرومانية والسويدية. وقد أنتجت الشبكة العالمية الدولية للتحقق من الوقائع Factchecking Network لعبة بطاقات لتأدية الأدوار تستهدف المراهقين الأكبر سناً (15–18 سنة) (تتوفر حالياً بالإنكليزية والإيطالية والبرتغالية والإسبانية)[379] يؤدي في إطارها الطلاب دور صحفيين في غرف أخبار يغطون استفتاءً مثيراً للجدل، تشوبه الدعاية عبر الإنترنت والمعلومات المُضلّلة. وقد جمع المعهد الأمريكي للصحافة أمثلة أخرى (Funke & Benkelman, 2019). على ذلك، اعترفت منظمة اليونسكو وحكومات مختارة بالعلاقة القائمة بين الدراية الإعلامية والمعلوماتية والألعاب خلال جلسة رائدة نُظّمت في فسحة ألعاب ريادية Games Bar في أواخر عام 2019 [380].

تتوفّر أيضاً مقاربات أكثر تقليدية تستند إلى المدرسة لتعزيز الدراية الإعلامية، وهي تستهدف مرحلة ما قبل سنّ المراهقة والمراهقين، إذ يبدأون عندها بالاهتمام بوسائل التواصل الاجتماعي والأخبار والسياسة. وقد سُجّلت استجابة بارزة بقيادة الحكومة في فنلندا، حيث يتضمّن نظام التعليم الرسمي فصولاً لتعزيز الدراية الإعلامية (Henley, 2020). وقد ذُكر أن هذا الأمر جعل المواطنين الفنلنديين على أتمّ الاستعداد للتعرف إلى الأكاذيب على الإنترنت. وفي أماكن أخرى من العالم، تعمل المنظمات الإعلامية ومجموعات المجتمع المدني على سد الفجوة التي تبقى بعد تقديم الحكومة هذه الخدمة في المدارس. وتشمل الأمثلة

....

375 https://aboutbadnews.com/

376 https://fakey.iuni.iu.edu/

377 https://www.bbc.co.uk/news/resources/idt-8760dd58-84f9-4c98-ade2-590562670096

378 http://project-youcheck.com/game-french/

379 https://factcheckingday.com/lesson-plan

380 https://en.unesco.org/news/media-and-information-literacy-joins-games-learning

النادي الإعلامي المدرسي الذي يديره المركز الإفريقي للدراية الإعلامية والمعلوماتية[381]، وهو منظمة غير حكومية؛ وأنشطة الدراية الإعلامية والمعلوماتية الموجهة نحو الطالب في منطقة الشرق الأوسط وشمال إفريقيا التي تضطلع بها أكاديمية بيروت للتربية الإعلامية والمعلوماتية[382]. ومن الأمثلة الأخرى على ذلك Lie Detectors[383] – وهي مبادرة غير ربحية في بلجيكا وألمانيا يزور في إطارها الصحفيون الفصول الدراسية للتفاعل مباشرةً مع التلاميذ وتثقيفهم بشأن الدراية الإخبارية وممارسات التدقيق في الأخبار. وتُسجّل مبادرة مماثلة في فرنسا يقودها صحفيون من لوموند Le Monde (Roucaute, 2017)، إلى جانب مبادرة Redes Cordiais التي يقودها صحفيون في البرازيل[384].

من جهتها، أطلقت هيئة الإذاعة البريطانية BBC، بمساعدة المجلس الثقافي البريطاني، مبادرة تتّسم بطابع عالمي أكبر وتقضي بتوفير موارد عالمية لتعزيز الدراية الإعلامية في المدارس من حول العالم (BBC,2020a). وفي أماكن أخرى، قاد سياسيون برامج لنشر الدراية الإعلامية في أوساط أطفال المدارس الثانوية، وهي مبادرة يمكن أن تساعد في بناء القدرة على الصمود في وجه المعلومات المُضلّلة (Troop, 2017). ومن الأمثلة الأخرى، مبادرة MediaWise التابعة لبوينتر Poynter[385]، التي قدمت تدريباً وجهاً لوجه وعبر الإنترنت في مجال الدراية الإعلامية والمعلوماتية لأكثر من خمسة ملايين مراهق ومواطن، مع تركيز خاص على المجتمعات التي تعاني من نقص الخدمات.

يقضي نشاط تكميلي لمقاربات الدراية الإعلامية والمعلوماتية في المدارس بتدريب المعلمين على الدراية الإعلامية والمعلوماتية لتمكينهم من تقديم تدريب ناجح في هذا المجال للطلاب في المدارس. ولدى اليونسكو الكثير من الموارد ذات الصلة، كما لديها عملية عالمية لمراجعة إطار المناهج الدراسية للمدرسين وتحديثها على ضوء التطورات الأخيرة، كانتشار المعلومات المُضلّلة[386]. إنه لأمر أساسي لأنه يتيح للتدريب على الدراية الإعلامية والمعلوماتية في المدارس أن يتوسّع ويصبح مستداماً. وتشمل الأمثلة على ذلك مشروع Educamidia البرازيلي[387] ومشروع YouCheck! الأوروبي[388]. أما في فرنسا، فيجري العمل مع معلمي المدارس للتصدي لموضوع المعلومات المُضلّلة أو "infox" باللغة الفرنسية من خلال مبادرة Savoir*Devenir وغيرها[389].

مقاربات تستهدف المُحتوى الإعلامي: تظهر الدراسات الحديثة أن الجيل الأكبر سناً (فوق 50 سنة) في بعض البلدان يتمتّع بقدرة تقلّ عن المتوسط على التعرّف إلى المعلومات الوقائية (Gottfried & Grieco, 2018) وتذكّر الادعاءات الكاذبة التي كُشفت بالفعل (Mantzarlis, 2017). ويتمثّل بالتالي أحد التحديات الهامة في تحديد أفضل السبل لتوفير الكفايات الإعلامية والمعلوماتية لهذه الفئة السكانية. فالمقاربات

381 https://www.africmil.org/programmes-and-projects/media-information-literacy/school-media-clubmedia-in-education

382 https://mdlab.lau.edu.lb

383 https://lie-detectors.org/

384 https://www.redescordiais.com.br/

385 https://www.poynter.org/mediawise/

386 https://en.unesco.org/sites/default/files/belgrade_recommendations_on_draft_global_standards_for_mil_curricula_guidelines_12_november.pdf

387 https://educamidia.org.br/habilidades

388 http://project-youcheck.com/about/

389 http://savoirdevenir.net/mediatropismes/

التي نوقشت سابقاً ليست مناسبة، إذ أن احتمال استخدام كبار السن الألعاب عبر الإنترنت أقلّ، وهم يعتمدون بشكل أقل بكثير على منصات التواصل الاجتماعي كمصدر للأخبار (Ofcom, 2018a). وتقضي إحدى الطرق الواعدة بتقديم برامج خاصة من خلال القنوات التلفزيونية الرئيسية. وفي إطار مشروع Beyond Fake News، قامت هيئة الإذاعة البريطانية BBC بتطوير سلسلة كاملة من الأفلام الوثائقية والتقارير الخاصة والمقالات الإخبارية عبر شبكاتها في أفريقيا والهند ومنطقة آسيا- المحيط الهادئ وأوروبا والأمريكيتين، وهي تُقدّم عبر التلفزيون والإذاعة والإنترنت (BBC, 2018c).

تدريب المؤثرين عبر الإنترنت والمنظمات الشبابية على الدراية الإعلامية والمعلوماتية: يستطيع المؤثرون الرقميون الذين لديهم الملايين من المتابعين نشر معلومات مضللة على نطاق واسع. وفي هذا الإطار، أطلق صحفيون في البرازيل مبادرات تدريبية مكرّسة للدراية الإعلامية والمعلوماتية هدفها تحسين قدرة هؤلاء المشاهير على التحقُّق من الوقائع في المُحتوى المنشور على الإنترنت قبل نشر منشورات تدعم المُحتوى المغلوط أو المضلل (Estarque, 2020). ويمثل التدريب على الدراية الإعلامية والمعلوماتية من خلال المنظمات الشبابية مقاربة واعدة أخرى يجري استكشافها في الهند وكينيا ونيجيريا[390]، بدعم من اليونسكو.

باتت مجموعات أدوات التدقيق والمواد التثقيفية عبر الإنترنت التي تسعى إلى تحسين فهم الجمهور للتدقيق والتحقُّق من الوقائع متاحة بشكل متزايد. ويمكن في هذا الصدد ذكر Edutopia[391]، وهي مخطط للدروس وضعته نيويورك تايمز (Schulten & Brown, 2017)، وحملة اليونسكو MIL CLICKS[392] والدورات الالكترونية المفتوحة الحاشدة MOOCs الخاصة بالمنظمة لتعزيز الدراية الإعلامية والمعلوماتية، والتفكير الناقد، والإبداع، والمواطنة، والمهارات ذات الصلة، مع مواد متوفرة بلغات متعددة (مثل: العربية والإنجليزية والفرنسية واليونانية والإسبانية). وبدعم من مكتب الأمم المتحدة المعني بالمخدرات والجريمة، جرى العمل في جنوب أفريقيا على تحديد منهاج نموذجي للأمم المتحدة بشأن المعلومات المُضلّلة والأخلاقيات[393].

في محاولة لتسهيل فهم عملية التدقيق في المُحتوى، أنتجت الشبكة الدولية للتحقق من الوقائع رسوماً متحركة من 7 خطوات للتحقق من الوقائع[394]، وهي متاحة حالياً باللغات الإنجليزية والفرنسية والإيطالية واليابانية والبرتغالية والصربية والسواحيلية. إلى جانب ذلك، أنتجت مؤسسة FullFact الخيرية المستقلة للتحقق من الوقائع في المملكة المتحدة مجموعة أدوات مماثلة من 10 خطوات للكشف عن المعلومات المغلوطة[395]، فضلاً عن مجموعة من المواد الموجهة للأطفال هدفها تعزيز الدراية. ويتمثل مثال آخر في فيديو تثقيفي عبر الإنترنت عنوانه «كيفية اكتشاف الأخبار الزائفة» أو How to spot fake news أنتجته منظمة FactCheck.org[396].

....

390 http://www.unesco.org/new/en/communication-and-information/resources/news-and-in-focus-articles/all-news/news/unesco_supported_mil_training_in_india_three_days_of_learni/

391 https://www.edutopia.org/blogs/tag/media-literacy

392 https://en.unesco.org/MILCLICKS

393 http://www.cfms.uct.ac.za/news/media-ethics-workshop-localizes-un-curriculum

394 https://factcheckingday.com/articles/24/this-cartoon-has-7-tips-for-fact-checking-online-information

395 https://fullfact.org/toolkit/

396 https://www.youtube.com/watch?reload=9&v=AkwWcHekMdo&feature=youtu.be

مخرجات تهدف إلى تحسين الاحتراف الصحفي

يتـم إنتـاج عـدد متزايـد مـن المـواد المرتبطـة بالإلمـام بالتدقيـق وبرامجـه الموجّهـة نحـو الصحفييـن، مثـل وحـدة التعلـم حـول تاريـخ التّضليـل (Posetti & Matthews, 2018). وفـي هـذا الصـدد، يُعتبـر دليـل اليونسـكو المعنـون «الصحافـة و«الأخبـار الزائفـة» والتّضليـل للتدريـس والتدريـب فـي مجـال الصحافـة» (Ireton & Posetti, 2018) مـورداً مجانيـاً يتوفّر بإحـدى عشـرة لغـة، مـع 30 ترجمـة إضافيـة كانـت معلقـة وقـت كتابـة هـذا التقريـر[397]. إلـى جانـب ذلـك، يوفـر مشـروع First Draft أيضـاً دورات للصحفييـن تتمحـور حـول التحقُّـق مـن المـواد الإعلاميـة والمواقـع الإلكترونيـة و«الميمـات» Memes المرئيـة ومقاطـع الفيديـو التـي تـم التلاعـب بهـا[398]. وقـد أطلـق المشـروع مؤخـراً صفحـة مـوارد خاصـة بجائحـة فيـروس كورونـا أيضـاً[399]. وتُسـتكمل هـذه المـوارد بالطبعـة الأخيـرة مـن «دليـل التدقيـق «Verification Handbook» (Silverman, 2020) الـذي يؤمّـن مشـورة حـول التحقيـق فـي المُحتـوى والمنصـات المتلاعـب بهـا وحمـلات التّضليـل.

يهـدف نـوع ثـانٍ مـن المـوارد المشـتركة للصحفييـن إلـى تعزيـز الدقـة فـي الإبـلاغ. ومـن الأمثلـة علـى ذلـك توفيـر مـورد جديـر بالثقـة لتنظيـم مُحتـوى أحـدث الأبحـاث حـول مواضيـع إخباريـة رئيسـية[400]، ومشـورة حاليـة بشـأن اسـتراتيجيات إشـراك وسـائل الإعـلام[401]، ومـورد مركـزي للبيانـات الحكوميـة العامـة[402] أو معلومـات وإحصـاءات تـمّ التحقُّق بدقـة مـن وقائعها تتعلّق بالاقتصـاد، والرعايـة الصحيـة، والهجـرة، ومـا إلـى ذلـك[403]. غيـر أن مـوارد كثيـرة مـن هـذا القبيـل تخصّ فـي الوقـت الحالـي بلـداً معيّنـاً، كانـت قـد صـدرت بلغـة محـددة، وهـي مُصمَّمـة للاسـتخدام البشـري اليـدوي. ويمكـن زيـادة فائدتهـا فـي مجـال التحقُّـق مـن الوقائـع والمُحتـوى إذا مـا أصبحـت أيضـاً قابلـة للقـراءة/يمكـن الوصـول إليهـا آليـاً، وفقـاً للمعاييـر المعمـول بهـا لتبـادل البيانـات.

باتـت جهـات متعـددة تـدرك فـي الوقـت الحالـي أن الصحفييـن قـد يسـتفيدون مـن أحـدث الأبحـاث الأكاديميـة عـن المعلومـات المُضلّلـة وأن باسـتطاعتهم حتـى البـدء فـي التعـاون مـع الباحثيـن مـن أجـل دمـج النتائـج التـي توصّلـت إليهـا أحـدث التطـورات فـي علـم النفـس والعلـوم الاجتماعيـة وعلـوم البيانـات (Lazer et al., 2017). وبالإمـكان أيضـاً التعلّـم مـن الخبـراء فـي مجـال تأثيـر المعلومـات والتواصـل الاسـتراتيجي (Jeangene Vilmer et al., 2018) بشـأن أفضـل الاسـتراتيجيات لكشـف المعلومـات المُضلّلـة، علـى سـبيل المثـال.

مـع تزايـد حجـم المعلومـات المُضلّلـة وتعقيدهـا، يتعيّـن علـى الصحفييـن أيضـاً أن يطّلعـوا علـى الاسـتخبارات الـواردة مـن مصـادر مفتوحـة[404] OSINT (Open Source Intelligence) التـي بـرزت مؤخـراً وعلـى أدوات التدقيـق فـي المُحتـوى وأفضـل الممارسـات لاسـتخدامها. وقـد بـدأت بعـض المنظمـات مشـاركة قوائـم الأدوات

397 https://en.unesco.org/fightfakenews
398 https://firstdraftnews.org/en/education/learn/
399 https://firstdraftnews.org/long-form-article/coronavirus-resources-for-reporters/
400 https://journalistsresource.org/
401 https://mediaengagement.org/
402 https://datausa.io/
403 https://fullfact.org/finder/
404 https://en.wikipedia.org/wiki/Open-source_intelligence

المُوصى بها وكيفية استخدامها، مثل: مجموعة أدوات First Draft[405]، وBusinessWorld[406] في الهند. كما بدأت الأدوات المتخصصة المُستخدمة على نطاق واسع (مثل: البرنامج المساعد للتدقيق InVID/ WeVerify[407]) بإنتاج مقاطع فيديو تثقيفية وتوثيق عبر الإنترنت لتمكين الصحفيين من تعلّم التقنيات واعتمادها بسرعة في عملهم. وقد حال نقص التمويل دون إتاحة هذه المواد للصحفيين بلغات متعددة.

4.2.7 دراسة حالة بشأن الاستجابة: معلومات مضللة حول جائحة فيروس كورونا

في سياق «مشكلة التّضليل» المرتبطة بجائحة فيروس كورونا، يتم تنفيذ الكثير من التدابير التثقيفية رقمياً – وغالباً ما تُستخدم البيئات نفسها عبر الإنترنت التي تتكاثر فيها المعلومات المُضلّلة (مثل: وسائل التواصل الاجتماعي). وتقوم مشروعات الدراية الإعلامية والمعلوماتية، ووسائل الإعلام، ومنظمات المجتمع المدني الموجهة نحو الصحافة، ومدارس الصحافة، والحكومات بنشر هذه الاستجابات في جميع أنحاء العالم.

تشمل الأمثلة على مشروعات الدراية الإعلامية والمعلوماتية ما يلي:

- نشرت صحيفة Dawn الباكستانية دليلاً مقتضباً للمواطنين للنأي بأنفسهم من الوقوع في فِخاخ مشكلة التّضليل في إطار جهد مبذول لتعزيز الدراية الإعلامية الرقمية (Jahangir, 2020).
- نشرت كلية لندن للاقتصاد دليلًا لمساعدة الأطفال على تصفّح الإنترنت والإبحار بعيداً عن المعلومات المُضلّلة بشأن جائحة فيروس كورونا تتوجّه فيه إلى العائلات التي أجبرتها الجائحة على تعليم أطفالها في المنزل (Livingstone, 2020).

تركّز التدخلات التثقيفية الموجهة للصحفيين على التدقيق والتحقُّق من الوقائع والإبلاغ عن مسائل الصحة بصورة أخلاقية. وترد في ما يلي بعض الأمثلة:

- تم تطوير دورة مجانية عبر الإنترنت[408] لتدريب الصحفيين على أفضل السبل لتغطية الجائحة من قبل مركز Knight للصحافة في الأمريكيّتين، بالشراكة مع منظمة الصحة العالمية واليونسكو، وبدعم من مؤسسة Knight وبرنامج الأمم المتحدة الإنمائي.
- تتضمّن صفحة الموارد بشأن جائحة فيروس كورونا التي وضعها مشروع First Draft «قاعدة بيانات لكشف الزيف»، وقائمة بالمصادر التي خضع محتواها للتنظيم، وندوات تثقيفية عبر الإنترنت حول الإبلاغ عن الوباء المتفشّي، وأدوات وأدلة للمساعدة في التدقيق في المعلومات الخاصة بجائحة فيروس كورونا وكشف زيفها.

....

405 https://firstdraftnews.org/long-form-article/coronavirus-tools-and-guides-for-journalists/

406 http://www.businessworld.in/article/5-Tools-Every-Young-Journalist-Should-Learn-About-To-Identify-Fake-News/01-04-2019-16868/

407 يوفّر برنامج التدقيق المساعِد InVID/WeVerify حالياً مقاطع فيديو تثقيفية على الانترنت، وصفّاً افتراضياً، ومساعدة تفاعلية: https://weverify.eu/verification-plugin/

408 https://www.ejta.eu/news/free-online-course-journalism-pandemic-covering-covid-19-now-and-future

- يستضيف المركز الأفريقي للتميّز الإعلامي (ACME) قائمة منظّمة المُحتوى تضمّ الموارد والأدوات والنصائح والمصادر المتصلة بالإبلاغ عن فيروس كورونا، بما في ذلك مجموعة للتدقيق في الوقائع.
- أنتجت المنظمة غير الحكومية الأفغانية NAI كتيباً بشأن «أساسيات الأداء الصحفي خلال فيروس كورونا»[409].
- أصدرت مجموعة الأبحاث Data and Society ورقة بها عشر نصائح للصحفيين الذين يغطون موضوع المعلومات المُضلّلة[410].

تكتسي المبادرات العابرة للحدود أهميةً خاصةً، لا سيما مثلاً «المنتدى العالمي للإبلاغ عن الأزمات الصحية العالمية» الذي أنشأه المركز الدولي للصحفيين (Barnathan, 2020) والذي يتضمن مركزاً تفاعلياً متعدد اللغات لآلاف الصحفيين حول العالم؛ هدفه المساعدة في الإبلاغ الأخلاقي المستنير من خلال الوصول المباشر إلى مصادر موثوقة للخبرة العلمية والطبية؛ وتسهيل تبادل المعرفة والتحقُّق من الوقائع/كشف الزيف بشكل تعاوني في مجال فيروس كورونا.

5.2.7 من هي الجهات الفاعلة الرئيسية وراء الاستجابات التثقيفية ومن يموّلها؟

شراكات بين أصحاب مصلحة متعددين: إنها مبادرات خاصة بالدراية الإعلامية والمعلوماتية تعمل في إطارها جهات فاعلة متعددة من فئات مختلفة ضمن شراكة. ومن الأمثلة على ذلك التحالف العالمي لليونسكو الخاص بالدراية الإعلامية والمعلوماتية GAPMIL [411] وشراكته مع تويتر خلال الأسبوع السنوي للدراية الإعلامية والمعلوماتية[412]، والعمل الذي تضطلع به الشراكة بشأن الذكاء الاصطناعي[413] في حقل الذكاء الاصطناعي ونزاهة الإعلام والذي تنخرط فيه أكثر من 100 منظمة، بما في ذلك جميع شركات ومنصات التواصل عبر الإنترنت الرئيسية، وبعض المؤسسات الإعلامية الرئيسية، ومراكز أبحاث ومنظمات غير ربحية. ومن الأمثلة الأخرى الحركة العالمية MisinfoCon [414] التي تُعنى على وجه التحديد بإنشاء أدوات للتدقيق والتحقُّق من الوقائع. وفي هذا الصدد، ينظّم مؤيدوها عروضاً لتجربة الأدوات، وهاكاثون (حدث واسع النطاق)، ومحادثات، ومناقشات تغطّي موضوعَي الدراية والتفكير الناقد.

مبادرات منظمات المجتمع المدني والمبادرات الشعبية: إنها برامج وموارد تتعلّق بالدراية الإعلامية والمعلوماتية تضعها منظمات لا تبغي الربح و/أو مواطنون. بالإضافة إلى الأمثلة التي وردت أعلاه (مثل: First Draft, Drog, LieDetectors)، تشمل أمثلة أخرى مبادرة Savoir*Devenir [415] المدعومة من «كرسي اليونسكو»؛

409 https://nai.org.af/law-and-legal-documents/

410 https://datasociety.net/wp-content/uploads/2020/04/10-Tips-pdf.pdf

411 https://en.unesco.org/themes/media-and-information-literacy/gapmil/about

412 https://en.unesco.org/news/unesco-partners-twitter-global-media-and-information-literacy-week-2018

413 https://www.partnershiponai.org/ai-and-media-integrity-steering-committee/

414 https://misinfocon.com/join-the-misinfocon-movement-f62172ccb1b

415 http://savoirdevenir.net/chaireunesco/objectifs-missions/

ومؤسسة 5Rights التي تركّز على إلمام الأطفال بالبيانات[416]؛ ومشروع Mafindo [417] الإندونيسي الشعبي لمُكافحة الخدع؛ ومركز الدراية الرقمية المموّل من غوغل[418] الذي يدرّب مدرسين وتلاميذ في جمهورية كوريا؛ وإشراك مجموعات الشباب في مشروعات تتمحور حول الدراية الإعلامية والمعلوماتية على مستوى أوروبا (مثل: INEDU [419])؛ والجهات الفاعلة الشعبية التي تنتج أشرطة فيديو وشروحات لكشف زيف المعلومات[420].

توفّر **المنظمات والشبكات** العاملة على التحقُّق من الوقائع أيضاً دورات تدريبية (موجّهة أساساً للصحفيين) وتنشر موارد تدريب، كمنظمات فردية أو من خلال مبادرات مشتركة[421]، ومن بينها الشبكات الدولية للتحقق من الوقائع (مثل الشبكة الدولية للتحقق من الوقائع[422]، وشبكة First Draft Partner [423]) والمنظمات الصحفية (مثل المركز الدولي للصحفيين[424]). وكثيراً ما تجتذب هذه المبادرات التمويل من شركات ومنصات التواصل عبر الإنترنت.

تُعتبر **المؤسسات الإعلامية** أيضاً ناشطة للغاية في تطوير مبادرات الدراية الإعلامية والمعلوماتية وتنفيذها، ليس فقط من خلال القنوات التقليدية (مثل: التلفاز) ووسائل التواصل الاجتماعي (مثل: YouTube)، ولكن أيضاً من خلال المشاركة المباشرة (على سبيل المثال: في الفصول الدراسية أو في ورش العمل والفعاليات التدريبية الموجهة نحو الصحافة). وقد نوقشت بعض الأمثلة في القسم 7.2.3 أعلاه. وتشمل أمثلة أخرى لتدريب الصحفيين والعمل التثقيفي الذي تضطلع به الشبكة الأفريقية لمراكز الإبلاغ الاستقصائي[425]، Code for Africa [426].

المبادرات الحكومية: باشرت حكومات كثيرة حالياً بجهود لتوفير الدراية الإعلامية والمعلوماتية المركّزة على التّضليل أو لدعمها. وتغطّي الأمثلة على هذه المبادرات (التي جمعت شبكة Poynter [427] الكثير منها) أستراليا، وبلجيكا، وكندا، والدنمارك، وفنلندا، وفرنسا، والهند[428]، وهولندا، ونيجيريا، وسنغافورة، والسويد، والولايات المتحدة. ومن بين الأمثلة، لابدّ من ذكر المبادرة التي أطلقها مركز الاتصال الفرنسي للتربية والإعلام[429] والتي تشمل المكتبات وأمناء المكتبات بشكل فريد بوصفهم من أصحاب المصلحة الرئيسيين في الاستجابة

....

416 https://5rightsfoundation.com/our-work/data-literacy/

417 https://www.mafindo.or.id/about/

418 https://www.blog.google/outreach-initiatives/google-org/digital-and-media-literacy-education-korea/

419 https://in-eduproject.eu/

420 تُدعى هذه السلسلة Smarter EveryDay، ويديرها مهندس يوتيوب Detin Sandlin:
www.youtube.com/watch?v=FY_NtO7SIrY، www.youtube.com/watch?v=V-1RhQ1uuQ4

421 Full Fact, Africa Check, Chequado: https://fullfact.org/blog/2020/feb/joint-research-fight-bad-information/

422 https://www.poynter.org/ifcn/

423 https://firstdraftnews.org/about/

424 https://www.icfj.org/our-work

425 https://investigativecenters.org/

426 https://medium.com/code-for-africa

427 https://www.poynter.org/ifcn/anti-misinformation-actions/

428 https://mgiep.unesco.org/

429 http://www.clemi.org/

المستندة إلى الدراية الإعلامية والمعلوماتية. من جهته، عمل الاتحاد الأوروبي بين عامَي 2016 و2018 على المستوى الأوروبي وقام بتمويل 10 مشروعات تتمحور حول الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية، علماً أن مشروعات أخرى هي قيد التفاوض؛ استناداً إلى نداء التمويل لعام 2019. وكانت غالبية هذه المشروعات تستهدف المواطنين (مثل YouCheck!)، فيما تستهدف البقية الصحفيين والإنتاج الإخباري (مثل: مجموعة الأدوات الأوروبية للدراية الإعلامية المخصّصة لغرف الأخبار).

شركات ومنصات التواصل عبر الإنترنت: تهدف المبادرات التثقيفية التي تضطلع بها هذه الشركات إلى:

- تعليم الأطفال مهارات الدراية الإعلامية والمعلوماتية، مثل مبادرة غوغل Be Internet Legends [430] ومبادرات يوتيوب ذات الصلة Be Internet Citizens [431]؛ ومبادرة غوغل العالمية Be Internet Awesome [432] (المُطبّقة حالياً بموارد محلية في الأرجنتين، وبلجيكا، والبرازيل، وتشيلي، وكولومبيا، وإيطاليا، والمكسيك، وبيرو، وبولندا، والمملكة العربية السعودية، والمملكة المتحدة، والولايات المتحدة)؛
- تدريب الصحفيين وتحسين درايتهم التكنولوجية ومهاراتهم، والاستثمار في المشروعات التحريرية الموجهة نحو الدراية الإعلامية، مثل برنامج Facebook Journalism Project [433] على فيس بوك، ومبادرة غوغل للأخبار Google News Initiative [434]، ومبادرة يوتيوب ذات الصلة[435]، وGoogle's Fact Check Explorer [436].

6.2.7 كيف يتم تقييم الاستجابات التثقيفية؟

يُعدّ تقييم نجاح مبادرات الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية في تغيير سلوك المواطنين لفائدة استخدام المعلومات المُضلّلة ومشاركتها مشكلة صعبة ومفتوحة إلى حد كبير. فوفقاً للأدلة التي استعرضتها الأبحاث لأغراض وضع هذه الدراسة، يبدو أن المقاييس المعيارية ومنهجيات التقييم لا تزال تفتقر إلى النضج. ويتمثل التحدي، على وجه الخصوص، في تجاوز مرحلة التوعية باتجاه تدخلات مستدامة ومؤسسية في مجال الدراية الإعلامية والمعلوماتية تؤدي إلى تغييرات دائمة وقابلة للقياس في سلوك المواطنين على الإنترنت.

وتبرز أيضاً الحاجة إلى إجراء تقييم مستقل لحياد المواد والتدريب في مجال الدراية الإعلامية والمعلوماتية وشموليتها، لا سيما تلك التي استحدثتها شركات التواصل عبر الإنترنت. وقد أثارت منظمات المجتمع المدني (5Rights Foundation, 2019) مخاوف من أنها تميل إلى التركيز على جعل المستخدمين (خاصةً الأطفال ولكن أيضاً الصحفيين) يركزون على المُحتوى الكاذب على حساب قضايا الخصوصية، بدلاً من الاستثمار في جهود ترمي إلى حل هذه المشكلات بأنفسهم. وقد أُشير أيضاً إلى أوجه قصور في الشمولية والشفافية

....

430 https://beinternetlegends.withgoogle.com/en-gb

431 https://internetcitizens.withyoutube.com

432 https://beinternetawesome.withgoogle.com/en_us

433 https://www.facebook.com/facebookmedia/solutions/facebook-journalism-project

434 https://newsinitiative.withgoogle.com/dnifund/report/european-innovation-supporting-quality-journalism/

435 https://youtube.googleblog.com/2018/07/building-better-news-experience-on.html

436 https://toolbox.google.com/factcheck/explorer

(5Rights Foundation, 2019)، نظراً للنقاش غير المناسب حول المخاطر الناشئة عن عملية تحديد السمات الخوارزمية، والمراقبة التلقائية للمُحتوى وتضخيمه، وتبعات الخصوصية المترتبة على جمع البيانات. وتوجد مخاوف مماثلة تتعلّق بالدعم الذي تؤمنه هذه الجهات الفاعلة التجارية للاستجابات التثقيفية الرامية إلى تعزيز الصحافة وتحسين مهارات الصحفيين كاستجابة للمعلومات المُضلّلة.

لابّد من إجراء المزيد من الأبحاث على نطاق واسع وعابر للمنصات لدراسة مسألة تعرض المواطنين للمعلومات المُضلّلة على الإنترنت ونشرها، وللنظر في الآثار المترتبة على فهم المواطنين لأنواع الاستجابات الأخرى للتضليل وتجاربهم معها. وفي هذا السياق، تكتسي مسألة قياس معرفة المواطنين بالاستجابات الخوارزمية واستجابات تنظيم المُحتوى الخاصة بالمنصات وفهمهم لها وطريقة تأثيرها على المعلومات المُضلّلة وحُريّة التّعبير والحق في الخصوصية والحق في الحصول على المعلومات أهميةً خاصة. ونظراً إلى الطبيعة الحديثة «للنيران المستعرة» للتضليل واسع النطاق على شبكة الإنترنت، لا تتوفّر مجموعة واسعة من الأبحاث التي تجيب على هذه الأسئلة الرئيسية، كما أن النتائج محدودة جغرافياً، ومتناقضة إلى حد ما في بعض الأحيان. وقد حفّز هذا الأمر واضعي السياسات والخبراء المستقلين على رفع توصيات إلى الحكومات للاستثمار في أبحاث إضافية حول هذه المسائل (HLEG report, 2018; DCMS report, 2018c; NED, 2018)، بما في ذلك الدراسات الإثنوغرافية وليس فقط مقاربة علم البيانات (NED, 2018).

وأثار الباحثون أيضاً مخاوف بشأن الاتجاه الأخير لتركيز الدراية الإعلامية والمعلوماتية في المقام الأول على التفكير الناقد بالأخبار (Frau-Meigs, 2019). وتكشف المخاوف، على وجه الخصوص، «أن هذا الأمر يجتذب الأموال التي يمكن أن تُخصّص، لولا ذلك، لمشروعات الدراية الإعلامية والمعلوماتية الكاملة؛ ففي إطار هذا الميل الأخير، يتمّ توفير تدخلات مدرسية لمرة واحدة دون متابعة كبيرة؛ كما أنّ التدخلات لا تنتشر على المستوى الوطني فتصل إلى عدد محدود من الطلاب».

7.2.7 التحديات والفرص

يتمثّل أحد التحديات في كيفية مساعدة الجمهور بشكل عام (لا سيّما أولئك الذين لديهم آراء مستقطبة) على فهم قيمة الدراية الإعلامية والمعلوماتية واستثمار الوقت لتعلم الانخراط بشكل واعٍ في وسائل التواصل الاجتماعي وممارسته. بالإضافة إلى ذلك، تصطدم الدراية الإعلامية والمعلوماتية بالقيود إذا لم تتجاوز التحقُّق من الوقائع والتدقيق في المُحتوى، وإن لم تشمل مهارات المواطنة الرقمية الأوسع نطاقاً- بما في ذلك حُريّة التّعبير وغيرها من الحريات على شبكة الإنترنت وخارجها (Frau-Meigs, 2019).

وفي ما يتعلق باستجابات الدراية الإعلامية والمعلوماتية واستجابات التعليم من أجل المواطنة العالمية التي تستهدف الأطفال، يكمن التحدي الرئيسي في تصميم مُحتوى وآليات تنفيذ تتّسم بقدر كاف من الإشراك وتتمتّع بتأثير دائم، نظراً إلى أنه يتعيّن على الاستجابات الموجهة نحو الأطفال، بحكم طبيعتها، أن تحقق نتائج متوسطة إلى طويلة الأجل. وهناك أيضاً تحد يتمثل في إدراج المبادرات الخاصة بالدراية الإعلامية والمعلوماتية التي تستهدف المعلومات المُضلّلة وتعزز التفكير الناقد في السياق الأوسع للتثقيف في مجال السلامة عبر الإنترنت. فعلى سبيل المثال، لابدّ من الربط بين مخاطر تصديق المعلومات المُضلّلة ومشاركتها من جهة، وما يتصل بذلك من مخاطر أوسع نطاقاً تتمثل في تحديد السمات، أو الإعلانات الموجهة بشكل دقيق، أو تقاسم موقع استناداً إلى نظام تحديد المواقع العالمي GPS من جهة أخرى.

وبالمثل، يجـب أن يتوفّر التدريب والتثقيـف الملائمـان بشـأن المخاطـر المهنيـة التـي يواجهها الصحفيـون الذيـن كثيـراً مـا تسـتهدفهم حمـلات التّضليـل. وتنتشـر هـذه الحمـلات عـادةً كراهيـة النسـاء والعنصريـة والتعصـب بصفتهـا تكتيـكات إسـاءة عبـر الإنترنـت تُصمّـم لتشـويه السـمعة والتّضليـل، وهـي تتطلـب إذاً اسـتجابات شـاملة تتنـاول السـلامة والأمـن الرقميَّيـن (Posetti, 2018a).

وهنـاك فئـة مسـتهدفة لا تركّـز عليهـا بصـورة خاصـة حمـلات الدرايـة الإعلاميـة والمعلوماتيـة، وهـي فئـة المواطنيـن الأكبـر سـناً، الذيـن تعتبرهـم بعض الأبحـاث أكثـر عرضـةً لتصديـق المعلومـات المُضلّلـة ونشـرها مـن الفئـات العمريـة الأخـرى (Vicol, 2020). وفـي الوقـت نفسـه، قـد يكـون اسـتخدام هـذه الفئـة للمنصـات الاجتماعيـة ومعرفتهـم بهـا وفهمهـم لهـا محـدودة للغايـة، ممـا يعقّـد التحـدي المتمثـل فـي إيجـاد أفضـل طريقـة لتصميـم حمـلات الدرايـة الإعلاميـة والمعلوماتيـة وتنفيذهـا بفعاليـة.

مـن التحديـات الأخـرى التـي يتعيـن التصـدي لهـا مـن خـلال المبـادرات التثقيفيـة؛ خلـق الوعـي بالأثـر السـلبي المحتمـل لاسـتخدام الميكنـة فـي المنصـات الإلكترونيـة، سـيّما وأن التقنيـات المميكنـة لمراقبـة المعلومـات المُضلّلـة التـي تُسـتخدم فـي بعـض البيئـات الإلكترونيـة قـد تُعانـي مـن التحيُّـز الخوارزمـي، وقـد تكـون تمييزيـة ضـد مجموعـة محـددة مـن المسـتخدمين (مثـل الفتيـات أو الجماعـات العرقيـة أو الإثنيـة). وقـد وجـدت أبحـاث حديثـة (5Rights Foundation, 2019) أنّ 83% مـن الذيـن تتـراوح أعمارهـم بيـن 11 و12 سـنة يؤيـدون قيـام المنصـات بإزالـة المُحتـوى آليـاً، دون الحاجـة إلـى أن يوسـمه مسـتخدم مـا. ولكـن لـم تتّضـح بعـد مـا هـي نسـبة هـؤلاء الأطفـال الذيـن يدركـون أيضـاً تبعـات الاسـتخدام غيـر المراقـب لهكـذا ميكنـة علـى حُريّـة التّعبيـر.

ويتعلق تحدٍ جديد بجائحة فيروس كورونا. وكما ذُكر سابقاً في هذه الدراسة، أشارت منظمة الصحة العالمية إلى وجود «أزمة معلومات»، أي مُحتوى مفرط يجعل من الصعب على الناس العثور على مصادر جديرة بالثقة وإرشادات موثوقة[437]. في سياق هذا المُحتوى الزائد حول الوباء، يكمن التحدي في تطوير قدرة الجمهور على التمييز بين المُحتوى الذي تم التحقُّق منه والمُحتوى المغلوط أو المضلل، وعلى التعرف إلى المُحتوى الذي هو في طور التقييم العلمي والمصادقة والذي لم يوضع بعد في أي من خانتَي المعلومات الصحيحة أو الكاذبة. ويقضي أحد التحديات ذات الصلة بالتنبّه إلى أن ردود الفعل التثقيفية إزاء خطر مشكلة التّضليل قصيرة الأمد بشكل حصري، مما يعني أنها تغفل الصلات المحتملة ببرامج التمكين المؤسسية طويلة الأمد وسياساته الرامية إلى تعزيز الدراية الإعلامية والمعلوماتية في مجال التّضليل بشكل عام، بما في ذلك لفائدة الأطفال وكبار السن.

أمـا بالنسـبة إلـى الفـرص، فقـد شـكّل الوبـاء أيضـاً نقطـة تركيـز جديـدة لوسـائل الإعـلام الإخباريـة والصحفييـن لإظهـار دورهـم المميـز وشـرحه. وقـد اسـتحدث هـذا الوبـاء أيضـاً لحظـة فريـدة مـن نوعهـا لتوعيـة المواطنيـن بحقـوق حُريّـة التّعبيـر وواجباتهـا، ولتثقيفهـم بغـرض مسـاعدتهم، ولتعزيـز المعرفـة والمهـارات المتصّلـة بالدرايـة الإعلاميـة والمعلوماتيـة والتعليـم مـن أجـل المواطنـة العالميـة. وتبـرز أيضـاً فرصـة لإطـلاق مبـادرات تثقيفيـة فوريـة هدفهـا التصـدي لمشـكلة التّضليـل، مـع مراعـاة التبعـات التثقيفيـة طويلـة الأجـل. ويمكـن تنظيـم هـذه المبـادرات بشـكل صريـح لضمـان نتائـج دائمـة فـي مجـال الدرايـة الإعلاميـة والمعلوماتيـة، بحيـث لا تتصـدى فحسـب للمعلومـات المُضلّلـة المتصلـة بجائحـة فيـروس كورونـا علـى وجـه الخصـوص بـل أيضـاً

[437] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf?sfvrsn=195f4010_6

لأنواع أخرى من المعلومات المُضلّلة الصحية والسياسية أو المناخية. وتوفر الجائحة إمكانيات للجمهور لتعلّم كيفية التعامل مع المُحتوى عبر النّظر والتشكيك فيه وليس من خلال التهكّم عليه، ولإصدار أحكام مستنيرة حول «مشكلة التّضليل» والاستجابات لها.

وفي الختام، يتمثّل تحدٍ هائل وفرصة رئيسية في الوقت نفسه، في جعل الدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية في متناول الأطفال في جميع أنحاء العالم، ذلك أن الأطفال يشكّلون بحسب التقديرات ثلث مستخدمي الإنترنت على الصعيد العالمي ويمثلون الجيل الذي سيتولى في الوقت المناسب المسؤولية عن المعلومات وغيرها من القضايا (Livingstone et al., 2016). لهذا السبب، ينبغي أن تدرج الحكومات في جميع أنحاء العالم الدراية الإعلامية والمعلوماتية في مناهجها المدرسية الوطنية، وأن تستثمر في التدريب المهني لمدرّسيها في مجال الدراية الإعلامية والمعلوماتية، وأن تعمل بشكل وثيق مع المجتمع المدني والمنظمات الإعلامية والجهات المستقلة المعنية بالتحقُّق من الوقائع وشركات ومنصات التواصل عبر الإنترنت من أجل ضمان توفير الدراية الإعلامية والمعلوماتية الشاملة والكاملة لجميع أصحاب المصلحة.

8.2.7 توصيات بشأن الاستجابات التثقيفية

تؤمّن التحديات والفرص المحددة أعلاه، وتبعاتها الهامة على حُريّة التّعبير، أساساً لخيارات العمل التالية في هذه الفئة من الاستجابات.

توصيات للمنظمات الدولية:

- العمل على تأمين مبادرات ومواد تثقيفية في مجال الدراية الإعلامية والمعلوماتية للبلدان والفئات السكانية التي لا يصلها في الوقت الراهن ما يكفي من هذه الخدمات، وباللغات غير المتوفّرة حالياً.
- التشجيع على اتباع مقاربة شاملة إزاء الدراية الإعلامية والمعلوماتية تغطّي قضايا حُريّة التّعبير، فضلاً عن المعلومات المُضلّلة في مختلف الموضوعات (مثل: الصحة والسياسة والبيئة).
- تشجيع الجهات المانحة على الاستثمار على وجه التحديد في التدابير المضادة للتضليل التي تعزز الدراية الإعلامية والمعلوماتية (بالإضافة إلى حُريّة التّعبير والصحافة المستقلة وتطوير الإعلام).

توصيات لشركات ومنصات التواصل عبر الإنترنت:

- دمج الدراية الإعلامية والمعلوماتية في استخدام خدماتها، وتمكين المستخدمين من فهم مجموعة القضايا ذات الصلة بالتّضليل، بما في ذلك التحقُّق من الوقائع والمسائل الخوارزمية ومسائل التصنيف.
- تعزيز المشروعات البحثية متعددة التخصصات المصمّمة لاختبار الاستجابات التثقيفية للتضليل، والإبلاغ عن هذه التجارب بطرق جادة تساعد على تقاسم المعرفة- عبر التخصصات الأكاديمية وبين هذا القطاع والمعلمين والباحثين.
- دعم تطوير استجابات عالمية و/أو إقليمية بشأن الدراية الإعلامية والمعلوماتية، لا سيما عن طريق تمويل المشروعات في المناطق التي لا تُغطيها حالياً هذه الخدمات بشكل كافٍ[438].

....

[438] انظر، على سبيل المثال، /https://www.ictworks.org/african-digital-literacy-project-grant-funding

توصيات للدول الفردية:

- وضع سياسات خاصة بالدراية الإعلامية والمعلوماتية وتخصيص الموارد لها أو تعزيزها، لا سيما في النظام التعليمي؛ حيث يحتاج المدرسون أيضاً إلى التدريب لإيصال الدراية الإعلامية والمعلوماتية إلى الأطفال والشباب كسبيل للتصدي للتضليل.
- تخصيص التمويل والدعم للتدخلات التي تستهدف المواطنين الأكبر سناً الذين هم في آن واحد فئة سكانية ناخبة رئيسية ومجموعة أساسية معرّضة لخطر نشر المعلومات المُضلّلة.
- دعم المبادرات الرامية إلى التصدي للتضليل الذي يستهدف الأطفال والشباب والنساء والمُتعطّلين عن العمل واللاجئين والمهاجرين والمجتمعات الريفية.

توصيات لوسائل الإعلام الإخبارية:

- استخدام منصاتها لتدريب الجماهير ومُستخدمي الإنترنت بشكل استباقي على التمييز بين المعلومات التي تم التحقُّق منها من جهة والمعلومات المُضلّلة من جهة أخرى، والمساعدة في تنمية المهارات المطلوبة للتعرف إلى ذلك والتنقل في النظام البيئي للمُحتوى الأوسع، مع مراعاة موضوع حُريّة التّعبير.
- دعم التدريب المتقدم على تقنيات التدقيق ومُكافحة التّضليل الاستقصائية لموظفي التحرير.
- التعاون مع مدارس الصحافة في مشروعات مُكافحة التّضليل التي يشارك فيها كل من الباحثين والطلاب لتحسين قدرات الخريجين وتعميق فهمهم وممارساتهم.

توصيات للمجتمع المدني:

- تعزيز الابتكار في مجال الدراية الإعلامية والمعلوماتية مثل الألعاب المضادة للتضليل، وتطوير طرق مبتكرة لتمكين الدوائر خارج النظام التعليمي المعرّضة لخطر التّضليل.
- دعم تطوير استجابات عالمية و/أو إقليمية بشأن الدراية الإعلامية والمعلوماتية، لا سيما في المناطق التي تعاني حالياً من نقص في الخدمات.
- إجراء تقييم مستقل لمبادرات الدراية الإعلامية والمعلوماتية التي تنفذها و/أو تدعمها شركات التواصل عبر الإنترنت.

توصيات للباحثين:

- وضع مقاييس لدراسة الدراية الإعلامية والمعلوماتية في إشكالية التّضليل وتطبيقها.
- التركيز على الأبحاث متعددة التخصصات لتطوير مقاربات تثقيفية جديدة تُعتبر تدبيراً مكافحاً للتضليل، مثل دمج الأساليب المستقاة من الدراسات الصحفية وعلوم الحاسوب وعلم النفس وعلم الاجتماع.. إلخ.
- إقامة شراكات مع منظمات إخبارية للمساعدة في تعزيز الإبلاغ الاستقصائي عن المعلومات المُضلّلة وتعميق نظرة الجمهور للتفاعل مع المُحتوى المضاد للتضليل.

3.7 استجابات التمكين وتصنيف المصداقية

المؤلفون: ديانا ماينارد، ودينيس تيسو، وسام غريغوري

تركّز الاستجابات التثقيفية على تعليم الناس أهمية التفكير الناقد والوعي الذاتي من حيث استخدام المعلومات، مما يعطيهم الكفايات العقلية الداخلية. من جهته، يبحث هذا القسم في الاستجابات التمكينية التي تركّز بشكل خاص على الأساليب والأدوات الخارجية وعلى المواقع التي تساعد المستخدمين على فهم طبيعة المعلومات ومصادرها بشكل فعلي. وبالتالي، يمكن القول إنّ هذين النوعين من الاستجابات يسيران جنباً إلى جنب.

كما نوقش في القسم السابق، يُعدّ تثقيف الصحفيين والمواطنين على حد سواء في مجال الدراية الإعلامية والمعلوماتية من الاستجابات الهامة. فحتى لو تعذّر إحباط التّضليل بالكامل، يمكن التخفيف من انتشار المعلومات المُضلّلة وتأثيرها إذا تمكّن الناس من استخدام التفكير الناقد في استخدامهم للمواد الإعلامية والمعلومات. وتدعم هذه الكفاية القدرة على التعرّف بشكل فعّال إلى المعلومات المُضلّلة وإلى جاذبيتها والمصادر التي قد تروج لها. وبإمكان هذا الوعي أن يدفع بأولئك المعرّضين للمعلومات الزائفة إلى فهم دورهم في منع انتشارها وتأثيرها.

يأتي هذا القسم ليكمل الاستجابات التثقيفية من خلال دراسة الجهود المبذولة لاستخدام أدوات التدقيق في المُحتوى ومؤشرات المُحتوى الإلكتروني التي يمكن اعتبارها أدوات مساعِدة أو محفّزة لكفايات الناس في مواجهة المعلومات المُضلّلة. وتهدف هذه الأدوات والمؤشرات إلى مساعدة المواطنين والصحفيين على تجنب الوقوع فريسة للتضليل على الإنترنت، وإلى تشجيع الممارسات الجيدة بين الصحفيين وفي صفوف شركات ومنصات الإنترنت والإعلام عند نشر المعلومات.

ويتناول القسم أيضاً الجهود التي تبذلها وسائل الإعلام الإخبارية لتعزيز مصداقيتها مقارنةً بمصادر أقل موثوقية، من خلال تسليط الضوء على العلامات التجارية والبث الإذاعي العام الموثوق، فضلاً عن الأساليب التي تستهدف المستهلكين لتقييم مصداقية مصادر الأخبار وموثوقيتها ولتصنيفها. ومن الأمثلة على ذلك Newsguard[439]، وDecodex[440]، والمؤشر العالمي للتضليل Global Disinformation Index[441]، ومعايير الشفافية في مشروع Trust Project[442]، وعدد من ملحقات المتصفّح (التي نوقش الكثير منها في فصول أخرى من هذا التقرير البحثي والتي تُعتبر خارجية بالنسبة إلى استخدام المُحتوى عبر تطبيقات مثل فيس بوك). وقد نوقشت أيضاً مسألة ظهور تطبيقات «التقاط الصور الخاضع للسيطرة controlled capture» (مثل TruePic[443]) والمقاربات التي تم تطويرها حديثاً للمصادقة على المصدر وتتبّعه والتي يُنظر في إمكانية استخدامها من قبل الأفراد والكيانات الإعلامية على حد سواء، وهي تشمل الأعمال التي

---

439 https://www.newsguardtech.com/

440 https://www.lemonde.fr/verification/

441 https://disinformationindex.org/

442 https://thetrustproject.org/

443 https://truepic.com/

اضطلع بها مشروع News Provenance[444]، ومبادرة Content Authenticity Initiative[445]، ومبادرة Journalism Trust Initiative التكميلية[446].

تتوفّر أمثلة عن جهود الربط بين كفايات الدراية الإعلامية والمعلوماتية وتوفير الأدوات المساعدة للمستخدمين، كمجموعة الأدوات التثقيفية التي وضعها Full Fact[447]. يوفّر هذا المورد منهجيات واقتراحات لأدوات تساعد في فهم مصدر الأخبار، وما هو ناقص، والشعور الذي تولّده الأخبار في نفوس القرّاء. وتقدم هذه المبادرات، كتلك الوارد وصفها هنا، المساعدة للمستخدم لتنفيذ هذه الأفكار عن طريق جعل اكتشاف المصدر الأصلي لمعلومة ما ومدى مصداقيتها عملياً على سبيل المثال.

على عكس أنواع أدوات التحقُّق من الوقائع التي وُصفت في القسم 6.2 والتي تحاول منع انتشار المعلومات المُضلّلة مباشرةً أو تحاول إعطاء إجابة محددة على السؤال التالي «هل هذه المعلومات صحيحة؟»، تلقي أدوات التمكين والمصداقية التي تُناقش هنا، بدلاً من ذلك، العبء على المستخدم ليفسّر الإشارات الصريحة المُعطاة عن المُحتوى. على سبيل المثال، توفّر الإشارات التي تقدمها مجموعات مثل: Credder وNewsguard (الموضحة أدناه) معلومات حول موثوقية المصدر أو مصداقيته، ولا تقدم إجابات عما إذا كانت معلومة ما صحيحة أم لا. وبالمثل، تُظهِر مبادرات تتبُّع المصدر المكان الذي نشأت فيه المعلومة، وتبيّن كيف نشأت في بعض الحالات، ولكنها تترك للمستخدم مهمّة تفسير ذلك (على سبيل المثال، من خلال فهم أنه إذا تم التقاط صورة قبل عدة سنوات، فقد لا تكون ذات صلة بحدث سُجّل للتو أو تم الإبلاغ عنه على هذا النحو). وبالتالي، تُعد هذه الأدوات مختلفة عن تلك التي تناولها الفصل السابق، كما سيرد بالتفصيل في القسم 7.3.3.

غالباً ما ترتبط الأدلة التي تثبت دقة مسألة ما ومصداقيتها بمعرفة الطرف الذي استحدث المُحتوى أصلاً أو شاركه أولاً. ويتعيّن على غرف الأخبار والأشخاص الذين يعتمدون على وسائل التواصل الاجتماعي للحصول على معلومات التحقيق في المصدر، قبل أن ينظروا في المُحتوى في حد ذاته. ويتعيّن مثلاً على الأشخاص إجراء أبحاث روتينية عن تاريخ معلومةٍ ما وموقعها متى أُدرجا فيها.

تماماً كما نوقش في مُختلف أقسام هذا التقرير، خاصةً في الفصل 3، كثيراً ما ترتبط المعلومات المُضلّلة بالحملات السياسية المحلية والأجنبية على حد سواء، ويمكن أن تؤدي إلى انعدام الثقة على نطاق واسع في سلطة الدولة. وتكمن إحدى الطرق التي تستطيع من خلالها الدول تبديد مخاوف الجمهور ومساعدته على التمييز بين هذه المعلومات وتلك الجديرة بالثقة –الأمر الذي يوفّر بالتالي ترياقاً لبعض المعلومات المُضلّلة– في زيادة الشفافية. ومما لا شكّ فيه أن إنشاء الدولة لنظام قوي للإفصاح الاستباقي، إلى جانب توفير المعلومات في الوقت الحقيقي، يُشكّلان معاً دعامة تعزّز وضوح المصدر. غير أنه يتعين التحلّي باليقظة؛ لأن المصدر لا يعني الوقائع ولا الشمولية. وعندما لا تكشف الدول بشكل كامل وفوري عن مجموعة الإحصاءات المتوفّرة حول جائحة فيروس كورونا مثلاً على القنوات التي يعرف الجميع أنها تخصّها؛ تفتح الباب أمام الشائعات والتكهنات المفهومة، ولكن أيضاً أمام مصادر لا أساس لها لسد الفجوة بمُحتوى مضلل.

....

444 https://www.newsprovenanceproject.com/

445 https://theblog.adobe.com/content-authenticity-initiative

446 https://rsf.org/en/news/rsf-and-its-partners-unveil-journalism-trust-initiative-combat-disinformation

447 https://fullfact.org/toolkit/

1.3.7 ما ومن الذي ترصده/تستهدفه استجابات التمكين وتصنيف المصداقية؟

تهدف مبادرات تتبُّع المصدر إلى مساعدة متلقّي الأخبار على فهم مصدر المعلومات، وبالتالي زيادة وعيهم بالمعلومات المُضلّلة، علماً أنّ هذه المبادرات تكمل الجهود التي تسعى إلى منع حدوثها ونشرها في المقام الأول. وتُعامَل هذه المبادرات معلومات الإسناد والبيانات الوصفية كأدوات يمكن أن تأتي بنظرة متعمّقة. والأمر ينطبق إجمالاً على الصور ومقاطع الفيديو الزائفة (مثل: الصور الزائفة الفجّة، ولكن أيضاً الصور التي أُسندت بشكل خاطىء) عن طريق مصادقة أفضل. وتشمل الأمثلة ما قام به تويتر، وTruepic، وSerelay، وAmber (ترد تفاصيل أكثر أدناه). وكبديل عن ذلك، قد تستهدف الطرق التي يتم بها عرض المعلومات على المستخدم، من أجل تنبيهه إلى معلومات مضللة محتملة، مثل المُحتوى القديم الذي أُعيد إظهاره للواجهة/ظهر كما ولو كان حالياً، والمصادر التي قد تثير الشكوك أو تكون غير جديرة بالثقة. ومن الأمثلة على ذلك مشروع News Provenance Project وأدوبي Adobe ، بالإضافة إلى سياسة تويتر بشأن التعامل مع المادة الإعلامية التي جرى التلاعب بها والتي توفر السياق داخل المنصة (Roth & Achuthan, 2020).

من ناحية أخرى، ترصد المبادرات القائمة على الثقة مزوّدي الأخبار أنفسهم في محاولة للتعرّف إلى الذين لا يمكن الاعتماد عليهم ولتقييم مقالات فردية ومؤلفين فرديين في بعض الأحيان. وتستهدف المبادرات القائمة على الاعتماد إلى حدٍّ كبير مزوّدي الأخبار، وتهدف إلى «إضفاء الشرعية» على أولئك الذين هم الأكثر جدارةً بالثقة. وتستهدف بعض المبادرات، مثل Newsguard، أيضاً مزودي وسائل التواصل الاجتماعي ووسائل الإعلام الوسيطة الأخرى، على أمل الحصول على عائدات مالية مقابل تزويد هذه الشركات بأدوات لتصنيف مؤمّني الأخبار ومصادرها.

2.3.7 مَن الذي تُحاول استجابات التمكين وتصنيف المصداقية مُساعدته؟

تستهدف هذه المبادرات خمسة أنواع رئيسية من الجماهير. ويحاول معظمها مساعدة عامة الجمهور من خلال تعزيز وعيه بالقضايا المحتملة، غير أنّ بعضها يستهدف أيضاً أولئك الذين يُنتجون الأخبار أو ينشرونها، مثل الصحفيين والمدوّنين، فضلاً عن المنظمات الإعلامية نفسها.

- يتم استهداف مُتلقّي الأخبار من خلال تزويدهم بأدوات تساعدهم على فهم طبيعة المُحتوى ومصادره بصورة أفضل (مثل: المصادقة على الصُّور ومقاطع الفيديو، والبيانات الوصفية الصريحة بشأنها، وعرض أفضل للمعلومات الزمنية، مثل: تواريخ النشر، وما إلى ذلك). ويشمل ذلك أيضاً تنبيه متلقّي الأخبار إلى الكيانات الإعلامية التي لا تفي بالمعايير المقبولة– كتلك التي لديها تاريخ من التمويل المشبوه أو من نشر مواد زائفة أو الاضطلاع بممارسات أخرى مشكوك فيها.
- يُستهدف مزوّدو الأخبار من خلال منحهم أساليب وأدوات يمكن تطبيقها على منصاتهم مباشرةً، من خلال، على سبيل المثال، إضافة البيانات الوصفية التي تشير إلى المصادر وغيرها من المعلومات المتعلقة بالمصداقية والتي تخصّ القصص والصور/مقاطع الفيديو التي يقدمونها، وكذلك عرض أفضل للمؤشرات المتصلة بالمصداقية.

- يُستهدف الصحفيون من خلال تزويدهم بأدوات تساعد على فهم طبيعة المقالات والمواد الإعلامية بشكل أفضل (على سبيل المثال، قضايا المصدر والمصداقية).
- يتم استهداف المدونين والصحفيين المواطنين والشركات الإعلامية من خلال توفير الممارسات الجيدة والمعايير التي يستطيع الجميع اتباعها عند إنتاج المواد (مثل مبادرة Journalism Trust Initiative).

تُستهدف شركات ومنصات التواصل عبر الإنترنت أيضاً من خلال أدوات مثل Newsguard. ويُنظر إليها على أنها سوق للخدمات التي يمكن أن تساعدها على التعرف إلى مروّجي المعلومات المُضلّلة، وخدمة جماهيرها باستخدام هذه الأدوات.

- تشير دراسة أجراها «مركز بيو Pew للأبحاث» (Mitchell, Gottfried et al., 2019) إلى أنّ مستهلكي الأخبار الأمريكيين يتوقّعون أن تعمد صناعة الأخبار إلى حلّ مشكلة التّضليل في الأخبار، على الرغم من أنهم يعتقدون أن السياسيين هم في المقام الأول مَن أحدثها. وتُلقي استجابات التمكين وتصنيف المصداقية بالعبء على المستخدم (وأحياناً على المنصة أو مؤسسات الإعلام الإخبارية) لانتقائية المُحتوى وتفسيره. وقد أشارت أبحاث مختلفة إلى القيمة المحتملة لشرح صحّة/عدم صحة خبر ما وتقديم معلومات وقائعية بديلة أو شرح مفصّل للسبب وراء المعلومات المغلوطة (Ecker & Lewandowsky, 2010; Swire & Ecker, 2018). ويسود افتراض مفاده أنه متى حصل الناس على مساعدة في هذا الإطار –تأتي على شكل إشارات عن الموثوقية– سيتلقون المعلومات المُضلّلة بمزيد من التشكيك.

تهدف هذه الاستجابات بالتالي إلى تيسير هذه المهمة بالنسبة إلى المستخدم النهائي عن طريق توفير آليات تدلّ على المعلومات المُضلّلة. ويمكن استخدام التغييرات في طريقة عرض الأخبار على المستخدمين، على سبيل المثال، لجعل الجمهور أكثر وعياً بالمُحتوى الذي يُحتمل أن يكون مُضللاً. ويتم، على وجه التحديد، اقتراح حلول لبروتوكولات يمكن من خلالها إضافة البيانات الوصفية الإعلامية أو جعلها أكثر وضوحاً، أو جعلها تُرافق المواد الإعلامية المنشورة أينما عُرضت (على سبيل المثال، عند زيادة مشاركتها على وسائل التواصل الاجتماعي أو في نتائج البحث على شبكة الإنترنت)، كما يرد في الأقسام التالية التي تصف مثل هذه المبادرات.

تستند المبادرات القائمة على الثقة، التي تسلّط الضوء على الممارسات الجيدة في وسائل الإعلام وتروّج للامتثال للمعايير المهنية وتكافئه، إلى فكرة مفادها أنه يمكن وسم مصادر إعلامية معينة على أنها جديرة بالثقة، مما يشجع على وضع معايير محددة ومعايير مرجعية مشتركة يستطيع أن يعتمدها كل الذين ينتجون المُحتوى الصحفي. وفي نهاية المطاف، يمهّد اعتماد هذه المعايير الطريق نحو عمليات التأكيد. ويمكن أن تفضي هكذا عمليات سواءً أكانت رسمية أم حتى غير رسمية إلى ممارسات تمييزية أفضل من جانب المستخدمين، وإلى ممارسات أفضل من قبل منتجي المواد الإعلامية. وتؤدي مبادرات الثقة إلى نتائج هامة، من بينها بناء ثقة المستخدمين في وسائل الإعلام والتصدي لمخاوفهم بشأن موثوقية المُحتوى.

يمكن تلخيص نظرية التغيير التي تجسّدها هذه المبادرات على النحو التالي:

- أسباب التّضليل ذات الصلة:
  - يقوم مزوّدو الأخبار أو شركات ومنصات التواصل عبر الإنترنت بنشر معلومات مضللة، إمّا لأنهم ليسوا جديرين بالثقة و/أو لأنهم لا يتعرّفون إليها؛
  - لا يتعرّف المستخدمون إلى المعلومات المُضلّلة، ويتأثرون بها وينشرونها بشكل إضافي أيضاً.

- الإجراءات التي تعتمدها المُبادرات لمعالجة الأسباب:
  - اعتماد مصادر أخبار جديرة بالثقة (وضمناً التشكيك في مصداقية تلك التي ليست كذلك)؛
  - تطوير أدوات تُمكّن وسائل الإعلام وشركات الإنترنت والمستخدمين خاصةً من اتخاذ قرارات أفضل بشأن المعلومات (والمصادر الإعلامية) التي يمكن الوثوق بها، فضلاً عن لفت انتباه الصحفيين والمحققين إلى مسائل معينة؛
  - وضع بروتوكولات لتوفير معلومات أفضل عن المصدر وتوعية المستخدمين بأهمية مصدر المُحتوى؛
  - تطوير تطبيقات «التقاط الصُّور الخاضع للسيطرة» controlled-capture applications تمكّن مبدعي الصور وموزعيها من نشر الثقة في محتواها.

- النتائج المرجوّة من المبادرات:
  - ممارسات تمييزية أفضل من جانب المستخدمين؛
  - ممارسات أفضل تعتمدها وسائل الإعلام الإخبارية وشركات الإنترنت؛
  - مُساءلة المصادر الإعلامية غير الجديرة بالثقة؛
  - تسهيل عملية اكتشاف المواد الإعلامية (الفيديو والصور) الزائفة وجعل نشرها أكثر صعوبةً؛
  - تحسين فهم الجمهور للتضليل وقواعده، وما إلى ذلك.

- الأثر المحتمل للمبادرات:
  - الثقة في وسائل الإعلام والتصدي للمخاوف بشأن موثوقية المعلومات، ممّا يؤدي إلى تحسين الثقة بدلاً من النسبية التهكّمية؛
  - تقليص الأُسُس المنطقية التي يستند إليها منتجو المعلومات المُضلّلة وتشجيع وسائل الإعلام متدنية المستوى على تعزيز جدارتها بالثقة؛
  - زيادة انتشار المعلومات الدقيقة والحدّ من انتشار المعلومات غير الدقيقة؛
  - تعزيز قُدرة المُبدعين/المُوزّعين على تأكيد جدارتهم بالثقة، وتمكين المستخدمين من تقييمها.

3.3.7 ما هي المخرجات التي تنشرها استجابات التمكين وتصنيف المصداقية؟

تنشر هذه المبادرات عدداً من المخرجات مختلفة الأنواع التي تهدف إلى مساعدة الجهات الفاعلة، وهي تتراوح ما بين المعلومات العامة والأساليب والبروتوكولات، بل وحتى الأدوات الفعلية في بعض الأحيان. ويمكن تلخيصها على النحو التالي:

- معلومات عن مصدر المواد والبروتوكولات التي تتيح توفيرها؛
- أدوات وموارد لتقييم مصداقية مصادر الأخبار؛ تصبّ في مخططات الاعتماد ونظم مراقبة المُحتوى؛
- أساليب وبروتوكولات لتحسين توفير المعلومات للمستخدم النهائي، ممّا يتيح زيادة الوعي بالمعلومات والمصادر الجديرة بالثقة وتلك غير الجديرة بالثقة؛
- أدوات لتصنيف مصادر الأخبار والمقالات والمؤلفين – ينفّذها مهنيُّون مُدربون (على سبيل المثال: Newsguard) أو يقودها المجتمع (مثل: Credder).

سيتمّ النظر في كل من هذه المبادرات بمزيد من التفصيل أدناه.

1.3.3.7 مبادرات تتبُّع المصدر

تنبثق مبادرات تتبّع المصدر من عدد من المصادر، ويمكن تقسيمها إلى ثلاث مجموعات فرعية رئيسية: أدوات وضعها مزوّدو الأخبار، وأدوات تُستخدم عند التقاط الصور/تصوير مقاطع الفيديو، واستجابات المنصات.

(1) تهدف **الأدوات التي يضعها مزوّدو الأخبار** إلى مُساعدة متلقّي الأخبار على التنبّه بصورة أكبر إلى المعلومات المُضلّلة، بدلاً من محاولة منعها من البروز في المقام الأول. على سبيل المثال، يهدف مشروع News Provenance إلى مساعدة المستخدمين على فهم طبيعة المُحتوى ومصادره بشكل أفضل (مثل: المصادقة على الصور ومقاطع الفيديو، والبيانات الوصفية الصريحة حولها، وتحسين عرض المعلومات الزمنية مثل تواريخ النشر، وما إلى ذلك).

(2) تهدف الأدوات التي تُستخدم عند **التقاط الصُور/تصوير الفيديو** إلى تتبّع البيانات الوصفية المُحسّنة والمصدر المعزّز، وتأكيد ما إذا كانت الصُور ومقاطع الفيديو قد عُدّلت أم لا. وفي هذا الإطار، تُعتبر TruePic[448] مثلاً شركة ناشئة مدعومة برأس مال استثماري تخطّط للعمل مع الشركات المصنّعة للأجهزة (تعمل حالياً مع Qualcomm فقط) لتسجيل الصور ومقاطع الفيديو لحظة التقاطها. من جهته، مكّن Serelay Trusted Media Capture[449] الهواتف المحمولة من التقاط الصور ومقاطع الفيديو التي يمكن

448 https://truepic.com/

449 https://www.serelay.com/

التحقُّق منها، كما أتاح للتطبيقات الأخرى الاستعلام عن الأصالة في وقتٍ لاحق. أما Amber [450]، فينتج أداتين هما Amber Authenticate لتسجيل بصمات الأصابع عند مصدرها وتتبع المصدر حتى التشغيل، وAmber Detect الذي يستخدم معالجة الإشارات والذكاء الاصطناعي لتحديد ملفات الصوت والفيديو المعدّلة. من جهته، يُعتبر Eyewitness to Atrocities [451] تطبيقاً للكاميرات المحمولة؛ تم تطويره لغرض محدد وهو توثيق الجرائم الدولية بحيث يمكن المصادقة على اللقطات لاستخدامها في التحقيقات أو المحاكمات. وعلى غِرار التطبيقات الأخرى، يقوم بتسجيل البيانات الوصفية عن وقت التسجيل ومكانه وبتخزينها آلياً، ويتضمّن سلسلة حيازة يمكن تعقبها. تُناقَش جميع هذه الأدوات (وغيرها) بمزيد من التفصيل في تقرير Witness (Witness Media Lab, 2019).

(3) تأتي استجابات المنصات مباشرةً من شركات ومنصات التواصل عبر الإنترنت، لا سيمّا منصات الصور والفيديو. ويقوم البعض منها بتشجيع المستخدمين على إضافة معلومات توضح أن المُحتوى يتوافق مع معايير الشركة ويجب عدم إزالته. ويسلّط يوتيوب الضوء على أهمية إضافة السياق[452]، لشرح أهمية الصور البيانية في مقاطع فيديو ذات أهمية إخبارية، على سبيل المثال (وبالتالي لتفادي رفضها آلياً من قبل يوتيوب في حال تم وسمها بسبب الشكّ فيها). وتقدّم الشركة مثالاً على تعليق صوتي عن تاريخ إحدى التظاهرات –علماً أنّ هذا النوع من المعلومات يساعد المستخدم على فهم مصدر شريط فيديو ما.

تتضمّن أنواع أخرى من الاستجابات تصنيف المُحتوى بصورة مباشرة؛ فعلى سبيل المثال، يستخدم يوتيوب هذه الطريقة لتوفير معلومات عن مقاطع فيديو تسلّط الضوء على نظريات المؤامرة (على سبيل المثال: حول هبوط أبولّو على القمر في عام 1969)؛ أو للإشارة إلى أنّ المُحتوى يصدر عن هيئة إذاعية تموّلها الدولة. وتتّخذ منصات أخرى إجراءات مماثلة بشأن معلومات غير دقيقة عن اللقاح. وقد عمد واتساب، في آب/أغسطس 2020، إلى اعتماد ميزة تشير إلى الرسائل التي تم إرسالها خمس مرات أو أكثر، كمؤشر عن معلومات محتملة واسعة الانتشار[453]. ويتيح النقر على رمز العدسة المكبّرة الذي يظهر تلقائياً بجوار رسالة من هذا النوع إطلاق عملية فحص عبر الإنترنت تهدف إلى الكشف عن أي نظرية مؤامرة معروفة أو معلومات مضللة مرتبطة بمُحتوى تلك الرسالة.

في أوائل عام 2018، بدأ يوتيوب بتصنيف المُحتوى لتحديد ما إذا كان المصدر يُعتبر «وسيلة إعلامية تموّلها الدولة» بحسب تعريف الشركة[454]. وفي حزيران/يونيو 2020، اعتمد فيس بوك سياسة مُماثلة، وأوضح أن هدفها مساعدة الناس على فهم ما إذا كانت الأخبار التي يقرأونها «تأتي من منشورة قد تكون تحت تأثير حكومة ما»[455]. واعتمد تويتر، بدوره، هذه الممارسة بعد بضعة أشهر[456].

....

450 https://ambervideo.co/

451 https://www.eyewitness.global/

452 https://support.google.com/youtube/answer/6345162?hl=en

453 https://www.theguardian.com/technology/2020/aug/04/whatsapp-launches-factcheck-feature-aimed-at-viral-messages

454 https://money.cnn.com/2018/02/02/media/youtube-state-funded-media-label/index.html

455 https://about.fb.com/news/2020/06/labeling-state-controlled-media/

456 https://www.bbc.co.uk/news/technology-53681021

2.3.3.7 المبادرات القائمة على الثقة والاعتماد

تهدف المبادرات القائمة على الثقة والاعتماد إلى وضع مجموعة متّفق عليها من معايير الثقة والشفافية للمصادر الإعلامية وتنفيذها. وتشمل هذه المعايير شفافية ملكية وسائل الإعلام ومصادر الإيرادات، فضلاً عن الأساليب الصحفية والامتثال للمعايير الأخلاقية والاستقلالية. ويهدف بعضها إلى التوصل إلى نظام اعتماد رسمي. وتشمل الأمثلة ما يلي:

- مبادرة Journalism Trust Initiative [457] (التي تضم منظمة «مراسلون بلا حدود» وشريكيها وكالة الصحافة الفرنسية واتحاد الإذاعة الأوروبي)؛
- مشروع Trust Project [458] (وهو اتحاد يضمّ أهم شركات الأخبار، بما في ذلك وكالة الأنباء الألمانية DPA، وThe Economist، وThe Globe and Mail، وتلفاز Hearst Television، و Independent Journal Review، وHaymarket Media، ومعهد Institute for Nonprofit News، و La Repubblica و La Stampa الإيطاليتَين، و Reach Plc، وواشنطن بوست، ويحظى بدعم خارجي من مختلف شركات التواصل الاجتماعي ومحرّكات البحث)؛
- مبادرة Trusted News [459] التي أنشأتها هيئة الإذاعة البريطانية BBC والتي تخطّط لعدد من الإجراءات التعاونية، مثل نظام الإنذار المبكر للاستجابة السريعة بحيث تستطيع المنظمات الإعلامية (وغيرها) تنبيه بعضها البعض بسرعة في حال كانت المعلومات المُضلّلة تهدّد حياة الإنسان. والمبادرة هذه مصمّمة خصيصاً لمنع تعطيل الديمقراطية أثناء الانتخابات، مع اتخاذ إجراءات أخرى تتمحور على وجه التحديد حول إعلام الناخبين وتثقيف وسائل الإعلام.

تنطوي المبادرات القائمة على الثقة أيضاً على استحداث أدوات وآليات لا تتيح للمستخدمين تقييم المصادر فحسب، بل أيضاً مقالات فردية و/أو صحفيين فرديين من حيث مصداقيتهم وجدارتهم بالثقة في بعض الحالات. ويعتبر **Credder** [460] مثلاً، الذي يصف نفسه على أنه «موقع موثوق لمراجعة وسائل الإعلام الإخبارية»، أنه ينبغي على «الأخبار أن تتنافس على الثقة، وليس على عدد النقرات». وهو يتيح للصحفيين والجمهور مراجعة المقالات، وقياس الثقة ليس فقط في المقالات نفسها ولكن أيضاً في المصادر المذكورة وفي المؤلفين، ويجمع إحصاءات عن هذه المسائل. وبشكل أكثر عموماً، تستخدم هذه الأدوات تقييمات ودرجات معطاة لجودة المصدر (استناداً إلى مقاييس مثل الدقة والموضوعية) لتوجيه المستخدمين نحو معلومات أعلى جودة ومساعدتهم في التنبّه إلى المعلومات متدنية الجودة وتجاهلها بشكل أفضل.

يوفّر **Newsguard** [461] برنامجاً مُساعِداً للمتصفّح يهدف إلى تقييم المواقع الإخبارية على أساس ما يعتبره ممارسات صحفية جيدة، من خلال منهجية تؤمن معلومات حيوية عند تحليل البيانات nutrition-label methodology. يتيح هذا الأمر تزويد القارىء بسياق إضافي لأخباره، كما يُحذّر المعلنين الذين قد يقلقون بشأن سمعة علامتهم التجارية لتجنب المواقع غير الموثوق بها. وتشير المواقع ذات التصنيف الأخضر

457 https://jti-rsf.org/

458 https://thetrustproject.org/

459 https://www.bbc.co.uk/mediacentre/latestnews/2019/disinformation

460 http://credder.com

461 https://www.newsguardtech.com/

إلى الممارسات الجيدة التي تتبع المعايير الأساسية للدقة والمُساءلة، في حين تُشير المواقع الحمراء إلى تلك التي لديها أجندة خفية أو التي تنشر عن علم؛ الأكاذيب أو الدعاية. بالإضافة إلى ذلك، تُشير المواقع الرمادية إلى منصات الإنترنت، في حين تشير المواقع البرتقالية إلى المواقع التهكّمية. ويرد رمز مشفّر بالألوان إلى جانب الروابط الإخبارية على محركات البحث وشريط الأحداث الجارية على وسائل التواصل الاجتماعي لإبلاغ الناس حتى قبل النقر على الرابط. ويمكن الحصول على معلومات إضافية حول الموقع -كسبب حصوله على التصنيف- عن طريق تحريك المؤشر mouse فوق الرمز و/أو النقر على زر للحصول على معلومات إضافية.

Décodex [462] أداة أنشأتها صحيفة لوموند Le Monde الفرنسية لمساعدة الناس على التحقُّق من المعلومات والتنبّه إلى الشائعات والأمور المبالغ فيها والحقائق الملتوية.. إلخ. تعمل الأداة بطريقتين: أداة بحث تمكّن المستخدم من التحقُّق من عنوان موقع ما لمزيد من المعلومات (مثلًا، لمعرفة ما إذا كان الموقع مصنفاً كموقع تهكّمي)؛ وملحق للمتصفّح يحذّر المستخدم عند انتقاله إلى موقع إلكتروني أو حساب على وسائل التواصل الاجتماعي شارك في نشر معلومات مضللة.

MediaBiasFactCheck [463] هي أداة تمكّن المستخدمين من التحقُّق من التحيز السياسي لمصدر إعلامي معيّن. ففي الولايات المتحدة على وجه الخصوص، يعتبر الجمهور أن التحيز واستخدام العبارات الغامضة أو المثيرة والأجندات هي الأسباب الأكثر شيوعاً لامتناع المواقع الإعلامية عن التمييز بوضوح بين الواقع والخيال، وفقاً لتقرير صدر عن معهد رويترز (Newman & Fletcher, 2017). واستناداً إلى الفلسفة الكامنة وراء هذه الأداة، من المفترض أن تكون المواقع الأقل تحيزاً هي الأكثر مصداقيةً، بفضل التقارير والمصادر الوقائعية المؤمّنة. أما المصادر المشكوك فيها، من ناحية أخرى، فتظهر ميزات مثل التحيز الشديد، واستخدام الكلمات المبطّنة (التي تحمل عاطفة قوية هدفها التأثير على القارئ)، وتعزيز الدعاية، وضعف الاستعانة بمواقع موثوقة كمصادر، والافتقار العام إلى الشفافية. ومع ذلك، لا تُعتبر المنهجية الكامنة وراء هذه المقاربة شفافة، وقد تعرّضت جودتها للانتقاد[464]. علاوةً على ذلك، لا يتّضح ما إذا كان من الجيد أن يخبرنا طرف واحد عن مدى تحيز مصدر إخباري، لأنّ الوضع غالباً ما يكون أكثر تعقيداً، ولابدّ لأيّ فكرة عن التحيز أن تستند إلى خط أساس.

Maldito Bulo [465] هو ملحق للمتصفّح أنشأه أحد المواقع الرئيسية لكشف زيف المعلومات في إسبانيا، وهو Maldita.es [466]. يحذّر البرنامج المساعِد؛ المستخدم الذي قام بتنزيله إذا كان الموقع الإلكتروني الذي يستشيره قد نشر معلومات مضللة، ويُطلعه على عدد القصص التي تم كشفها في اسم النطاق.

KnowNews [467] هو ملحق للمتصفح يهدف إلى مساعدة المستخدمين على التعرّف إلى المواقع الإخبارية الجديرة بالثقة أو الموثوقة. وقد طوّرته Media Monitoring Africa، وهي منظمة مستقلة لا تبغي الربح من

....

462 https://www.lemonde.fr/verification/

463 https://mediabiasfactcheck.com/

464 https://rationalwiki.org/wiki/Media_Bias/Fact_Check

465 https://chrome.google.com/webstore/detail/maldito-bulo/bpancimhkhejliinianojlkbbajehfdl

466 https://maldita.es/

467 https://newstools.co.za/page/knownews

جنـوب أفريقيـا تـروّج لحُريّـة الإعـلام وجودتـه وللصحافـة الأخلاقيـة[468]. ويقـوم ملحـق المتصفّـح هـذا بتصنيـف المواقـع الإخباريـة آليـاً علـى أسـاس مصداقيتهـا، فيعتبـر أن المواقـع موثوقـة، أو «مُراوِغـة»، أو غيـر مصنفـة. وتركـز الأداة علـى المُحتوى نفسـه، غيـر أنها تقيّـم أيضـاً المعلومـات مباشـرةً، مثـل أصالـة الصـورة وأساسـها. وقـد تـم تطويرهـا بالشـراكة مـع فيـس بـوك وغوغـل، فضـلاً عـن عـدد مـن المنظمـات الأخـرى.

أطلقـت **لجنـة نايت Knight** حـول الثقـة والإعـلام والديمقراطيـة، وهـي مجموعـة من الأشـخاص الذين يروجون لمجتمعـات أكثـر اسـتنارةً وانخراطـاً، مبـادرة الثقـة والإعـلام والديمقراطيـة Trust, Media and Democracy Initiative التابعـة لمؤسسـة The Knight Foundation[469]. تقـدم هـذه المجموعـة غيـر الحزبيـة التمويـل لسـبع مبـادرات هـي:

- Cortico لتحليلات البيانات لإبراز الأصوات الممثلة تمثيلاً ناقصاً؛
- Duke Tech & Check Cooperative + قاعدة بيانات Share the Facts؛
- شبكة First Draft للتحقق من الوقائع؛
- التحقُّق من الوقائع الخاص بـAP؛
- برنامج تدريب الصحفيين التابع لمعهد رينولدز للصحافة Reynolds Journalism Institute؛
- مؤشرات الثقة في مشروع Trust Project التابع لجامعة سانتا كلارا؛
- Your Voice Ohio لتعزيز العلاقات مع المجتمعات المحلية.

يتعـاون المجلـس الدولـي للاتصـالات الصحفيـة IPTC – وهـو هيئـة عالميـة لوضـع المعاييـر الخاصـة بوسـائل الإعـلام الإخباريـة – مـع الكثيـر مـن المبـادرات التـي تتمحـور حـول الثقـة والمعلومـات المغلوطـة فـي صناعـة الأخبـار منـذ عـام 2018. وينطـوي هـذا التعـاون أساسـاً علـى العمـل مـع مشـروع Trust Project ومبـادرة Journalism Trust Initiative التـي أطلقتهـا منظمـة «مُراسـلون بـلا حـدود»، ولكـن أيضـاً إلـى حـد مـا مـع Credibility Coalition، وCertified Content Coalition وغيرهمـا بهـدف تحديـد جميـع الوسـائل المعروفـة للتعبيـر عـن الثقـة فـي المُحتـوى الإخبـاري.

وفـي نيسـان/أبريل 2020، نشـر المجلـس الدولـي IPTC مُسـوّدةً لمجموعـة مـن المبـادئ التوجيهيـة[470] التـي تهـدف إلـى تمكيـن أي وكالـة إخبارية من إضافـة معلومـات الثقـة الخاصـة بهـا إلـى أيّ مـواد إخبارية توزّعهـا. ويمكـن أيضـاً تحويـل هـذه المؤشـرات إلـى لغـة ترميـز معياريـة schema.org تُضـاف إلـى صفحـات HTML وتُعالـج آليـاً بواسـطة محـركات البحـث ومنصـات التواصـل الاجتماعـي وأدوات مُتخصّصـة مثـل: البرنامـج المُسـاعِد NewsGuard. ويتيـح هـذا الأمـر للمسـتخدمين رؤيـة مؤشـرات الثقـة واتخـاذ قـرار بنفسـهم حـول جـدارة أي خبـر بالثقـة.

....

468 https://mediamonitoringafrica.org/

469 https://knightfoundation.org/topics/trust-media-and-democracy

470 https://iptc.org/news/public-draft-for-comment-expressing-trust-and-credibility-information-in-iptc-standards

تهدف المبادئ التوجيهية إلى تشجيع ناشري الأخبار على استخدام مؤشرات الثقة لإظهار السبب الذي ينبغي أن يدفع إلى الوثوق بهم، بدلاً من مجرد إظهار شهادة الجدارة بالثقة. وينبغي تشجيع القراء على متابعة الروابط لفهم القضايا على نحو أفضل. وتشمل المؤشرات تلك المتصلة بالسياسة التحريرية (مثل: البيانات المتعلقة بممارسات الإفصاح والتصويب، وتنوُّع المصادر وتسميتها، والأخلاقيات، وسياسات ردود الأفعال– Feedback)؛ والمؤشرات على مستوى الأطراف (مثل: قوائم الأعمال الأخرى التي وضعها المؤلف أو مقدم الخدمة، والجوائز التي فاز بها، ومجالات الخبرة)؛ والمؤشرات على مستوى المؤسسة (مثل تنوّع الموظفين، وتاريخ تأسيس المنظمة، وحجمها وما إلى ذلك)؛ والمؤشرات على مستوى العمل المحدد (مثل التفاصيل المتعلقة بالخط الزمني، والمحرّر، والتحقُّق من الوقائع، والتصحيحات، وموفر الخدمة)؛ والمؤشرات على مستوى الفرد (تفاصيل عن مؤلف المقال)؛ ومؤشر نوع العمل (على سبيل المثال: ما إذا كان تهكماً أم لا، ونوع التقرير، ومعلومات عن الخلفية، وما إلى ذلك).

تتبع هذه المبادئ التوجيهية فكرة تمكين المُستخدمين من خلال إتاحة المجال أمامهم لاتخاذ قراراتهم الخاصة بدلاً من اتباع ما يُقترح عليهم بشكل أعمى. وهي تسهّل أيضاً على منتجي المعلومات ومتلقّيها اتباع البروتوكولات المعمول بها. غير أنّ أحد العوامل المقيدة لهذا النوع من المنهجيات يكمن في أن المبادئ التوجيهية مُعقّدة للغاية، وتتطلّب الوقت والجهد من جانب المستخدم لتطوير فهم كامل عن الجدارة بالثقة، ومن ثم لتقييم مطابقتها لادعاءات المنظمة أو المُحتوى. ويناسب هذا الجهد الذهني الإضافي أولئك القادرين أصلاً على التمييز بطبيعتهم، وليس أولئك الأكثر عرضةً للتضليل.

وأخيراً، لابدّ من ذكر بعض المبادرات المحدّدة التي تركّز على نوع معين من الشائعات أو الموضوعات، مثل مشروع الثقة باللقاحات **Vaccine Confidence Project** (Larson, 2018). تركّز هذه المبادرة على الكشف المبكر عن الشائعات حول اللقاحات في محاولة لمنعها من اكتساب الزخم، ولكنها يدوية بشكل كامل. ويقوم فريق من الخبراء الدوليين برصد الأخبار ووسائل التواصل الاجتماعي، كما يبقي على مؤشر ثقة باللقاحات يستند إلى تتبّع المواقف العامة تجاه القضايا ذات الصلة. وفي حين يُعدّ هذا المشروع عملية للتحقق من الوقائع في المقام الأول، غير أنه يضطلع بأبحاث ذات صلة تتناول الثقة والمخاطر في هذا السياق، وهو مكرّس لبناء ثقة الجمهور والتخفيف من المخاطر في مجال الصحة العالمية. فمن خلال التنبّه إلى إشارات مبكرة عن عدم ثقة الجمهور ومن خلال التشكيك وتوفير مشورة وتحليل للمخاطر، يسعى المشروع إلى إشراك الجمهور في وقت مبكر فيمنع بالتالي الاضطرابات المحتملة في البرنامج.

4.3.7 مَن هي الجهات الفاعلة الرئيسية التي تقف وراء استجابات التمكين والمصداقية ومَن يموّلها؟

تتنوّع الجهات الفاعلة في هذه الأنواع من المبادرات ومصادر تمويلها بشكل كبير، فتتراوح بين وسائل الإعلام الإخبارية ووسائل التواصل الاجتماعي وشركات التواصل عبر الإنترنت، مروراً بمنظمات الرصد التي لا تبغي الربح.

عادةً ما يتم تمويل مبادرات الثقة والاعتماد إما من قِبَل الشركات الإعلامية التي تعمل معاً لتطوير أنظمة اعتماد رسمية، أو من قبل منظمات رصد مثل Media Monitor Africa ومبادرة Journalism Trust

Initiative. ومن الواضح أن شركات الإعلام لديها مصلحة في بناء الثقة في مصادر الأخبار، على الرغم من أن هذا الأمر يثير عدداً من المُعضلات الأخلاقية (انظر المزيد من المناقشة حول هذا الموضوع في القسم 7.3.7 أدناه).

تحظى مبادرات المصدر، هي أيضاً، بتمويل من مصادر متنوعة. وعادةً ما تموّل شركات التواصل الاجتماعي، مثل: يوتيوب وتويتر، استجابات المنصات، في حين يؤمّن مزودو الأخبار مثل مشروع News Provenance Project التابع لنيويورك تايمز أدوات أخرى. وغالباً ما تقوم شركات البرمجيات الخاصة بالصور والفيديو، كأدوبي Adobe، بتمويل أدوات تسُتخدم عند التقاط الصور/تصوير أشرطة فيديو، علماً أن هذه الأدوات قد تنشئها أيضاً شركات ناشئة مخصصة لذلك مثل TruePic [471] وAmber [472] وSerelay [473]. وتأتي التطبيقات مفتوحة المصدر إجمالاً من منظمات لا تبغي الربح، مثل Tella [474] التي تموّلها المنظمة غير الربحية Horizontal، وEyewitness to Atrocities التي تحظى بتمويل من «نقابة المحامين الدولية» في لندن بالشراكة مع LexisNexis. ويتم تمويل مشروع The Guardian Project، الذي ينتج التطبيق مفتوح المصدر ProofMode [475]، من قبل مجموعة متنوعة من المنظمات والمؤسسات، بما فيها Witness وغوغل وحكومات مختلفة.

5.3.7 دراسة حالة بشأن الاستجابة: معلومات مضللة حول جائحة فيروس كورونا

أبرزت جائحة فيروس كورونا ضرورة توفّر أساليب تساعد الجمهور على التحلّي بوعي أكبر بشأن المعلومات المُضلّلة. وفي حين تُشكّل الاستجابات التثقيفية نوعاً رئيسياً من التدخلات، حيث تنتج منظمات وحكومات كثيرة أدلة للاطّلاع على المعلومات والشائعات المشكوك فيها التي تُحيط بجائحة فيروس كورونا، تتوفّر حالات محددة قليلة قامت فيها المنظمات باستحداث تصنيف للمصداقية وآليات للتمكين أو بإلقاء الضوء عليها. وقد منحت شركات متعددة للتواصل عبر الإنترنت مكانة بارزة لمصادر المعلومات الموثوقة على خدماتها، لا سيّما مثلاً منظمة الصحة العالمية ووزارات الصحة الوطنية. وينطوي وضع المعالم Signposting عادةً على ذكر روابط لمصادر معلومات جديرة بالثقة، بدلاً من الإشارة صراحةً إلى مصادر غير جديرة بالثقة. ومن الأمثلة على هذه الجهود كلية الطب في جامعة هارفارد[476] التي تضع قائمة بمصادر موثوقة للمعلومات عن فيروس كورونا وتقدم نصائح حول اكتشاف هذا النوع من الموارد المعرفية.

هناك طريقة مثيرة للاهتمام تساعد في تسليط الضوء على المصداقية، اعتمدتها ويكيبيديا من خلال WikiProjectMedicine [477]، وهي مجموعة من حوالي 35,000 مقال يقوم برصدها حوالي 150 محرراً من الخبراء. وبمجرد وسم مقالة ما على أنها تتعلق بالطب، يتم فحصها عن كثب[478]. وبهذه الطريقة، تعمل

....

471 https://truepic.com/

472 https://ambervideo.co/

473 https://www.serelay.com/

474 https://hzontal.org/tella/

475 https://guardianproject.info/apps/org.witness.proofmode/

476 https://www.health.harvard.edu/blog/be-careful-where-you-get-your-news-about-coronavirus-2020020118801

477 https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine

478 https://www.wired.com/story/how-wikipedia-prevents-spread-coronavirus-misinformation/

ويكيبيديا كنموذج يُحتذى به من خلال اعتماد معايير منفصلة وقواعد أكثر صرامة لحالات معينة (في هذه الحال، الصحة العامة). وتبقي ويكيبيديا أيضاً على صفحة محدثة تتضمّن مجموعة متنوعة من «المعلومات المغلوطة» (كما تسمّيها) التي تتمحور تحديداً حول جائحة فيروس كورونا – 19 [479].

تُستكمل هذه الأمثلة بمبادرات ترمي إلى وسم المُحتوى المشكوك فيه، في محاولة لمساعدة الناس بشكل غير مباشر على فهم المصادر الحقيقية التي تتناول الجائحة، فيما توفّر أيضاً معلومات تم التحقُّق منها (عادةً ما تجري إزالة المُحتوى الذي يُعتبر أنه يشكّل خطراً على الصحة العامة وصحة الفرد حالما يتم التعرّف إليه). وتقدم مبادئ توجيهية متعددة إجمالاً نصائح للمستخدمين حول كيفية اكتشاف مصدر موثوق ولكن أيضاً مصدر غير موثوق. وغالباً ما توفّر المشورة بشأن المشاركة (كما تمت مناقشته أيضاً في القسم الأول من هذا الفصل المتعلق باستجابات الدراية الإعلامية والمعلوماتية).

من جهته، أطلق NewsGuard إحدى المبادرات الآيلة إلى وسم مصادر المعلومات المُضلّلة حول جائحة فيروس كورونا، بعد أن أقام شراكةً مع BT لإطلاق مجموعة أدوات عبر الإنترنت[480] هدفها تنمية الوعي بالبرنامج المساعِد للمتصفح التابع لـ NewsGuard بغية مساعدة الجمهور في المملكة المتحدة على تقييم أي معلومات تتعلق بالوباء العالمي يجدها على الإنترنت. وتحظى هذه المبادرة أيضاً بدعم من وزارة الثقافة والإعلام والرياضة في المملكة المتحدة وجمعية المكتبات في المملكة المتحدة. هذا وقد جعل NewsGuard البرنامج المساعِد للمتصفح الخاص به مجانياً حتى نهاية تموز/يوليو[481]، على ضوء جائحة فيروس كورونا بصورة خاصة. ولم يكن هذا البرنامج المساعِد في السابق متوفراً إلا كخدمة اشتراك، ولكن ليس بالنسبة إلى مستخدمي الأجهزة المحمولة لميكروسوفت–إدج Microsoft Edge. ومنذ 14 أيار/مايو 2020، وُضع بتصرّف جميع مستخدمي Microsoft Edge على كل من الهواتف المحمولة وتطبيقات الحاسوب المكتبي، شرط استخدام الملحق على هذا المتصفح وتحميله من متجر Microsoft Edge. كما تم إنشاء مركز لتتبّع المعلومات المُضلّلة حول جائحة فيروس كورونا[482] يشير إلى جميع مواقع الأخبار والمعلومات في الولايات المتحدة والمملكة المتحدة وفرنسا وإيطاليا وألمانيا التي تمّ تحديدها على أنها تنشر معلومات مغلوطة عن الفيروس[483]. وتشمل القائمة مواقع معروفة بنشرها مُحتوى صحي كاذب، فضلاً عن مواقع سياسية تتبنّى نظريات مؤامرة تتجاوز بكثير نطاق السياسة.

في حين ما زالت معظم استجابات تصنيف المصداقية المُستخدمة لمُكافحة المعلومات المُضلّلة في بداياتها بسبب القضايا الأخلاقية والقانونية والتكنولوجية التي لم تُحلّ بالكامل بعد، فإنّ الكم الهائل من المعلومات المُضلّلة حول فيروس كورونا وخطورتها المحتملة قد يتحوّل إلى قوة دفع باتجاه بذل المزيد من الجهود لتطوير أدوات تساعد المستخدمين على استكمال مستويات الدراية الإعلامية والمعلوماتية المتوفّرة لديهم وهم يُبحرون عبر «أزمة المعلومات».

....

479 https://en.wikipedia.org/wiki/Misinformation_related_to_the_2019%E2%80%9320_coronavirus_pandemic

480 https://www.newsguardtech.com/press/newsguard-partners-with-dcms-and-bt-to-help-counter-spread-of-covid-19-fake-news-as-misinformation-peaks/

481 https://www.zdnet.com/article/newsguard-drops-its-paywall-to-combat-coronavirus-information/

482 https://www.newsguardtech.com/coronavirus-misinformation-tracking-center/

483 في نهاية آذار/مارس 2020، كان عدد المواقع 144 موقعاً، على الرغم من أن هذا العدد يرتفع باستمرار.

6.3.7 كيف يتم تقييم استجابات التمكين وتصنيف المصداقية؟

إنّ الكثير من هذه المبادرات تعاونية إلى حد كبير في المنشأ والتطوّر، وهي بالتالي موجّهة بالجهود المجتمعية والمجالس الاستشارية وخاضعة لتقييمها.

- يجري تقييم المبادرات القائمة على المصدر على المستوى الداخلي إلى حد كبير. فعلى سبيل المثال، يجري مشروع News Provenance Project أبحاثاً عن المستخدمين لاختبار فعالية مقاربته المقترحة، في محاولةً لاكتشاف ما إذا كان تعزيز الوصول إلى البيانات الوصفية والمعلومات الداعمة يُساعد المستخدمين على فهم صحة الصحافة المصوّرة المنتجة مهنياً بشكل أفضل.

- تعمل المبادرات القائمة على الاعتماد على تطوير معايير مجتمعية ذات خلفية متينة. فعلى سبيل المثال، تشمل مبادرة Journalism Trust Initiative عدداً كبيراً للغاية من المنظمات الاستشارية، بما في ذلك هيئات رسمية لوضع المعايير، مثل: الرابطة الفرنسية لتوحيد المعايير (AFNOR) ونظيرها الألماني، المعهد الألماني لتوحيد المعايير (DIN)، فضلاً عن اللجنة الأوروبية لتوحيد المعايير (CEN).

- لا يتم دائماً تقييم المبادرات القائمة على الثقة بشكل رسمي، بل تعتمد على المُدخلات التي يوجّهها المجتمع المحلي. وتعرض مواقع مثل Credder، مثلاً، عدد المراجعات التي يقدمها المستخدمون (ومحتواها)، لذلك يسهُل استخلاص معلومات إحصائية حول استخدامها ومستويات الاتفاق. غير أنّ المسألة الأقل وضوحاً هي مدى فائدة هذه المراجعات بالنسبة إلى الآخرين. وبعبارةٍ أخرى، يمكن بسهولة الحكم على جودة المدخلات (المراجعات نفسها) ويمكن تقييم الثقة الجماعية، غير أنّ فهم فائدة الناتج وتأثيره العام أقل سهولةً. وتستخدم مواقع أخرى، مثل Newsguard، مُحللين مُدرَّبين هم صحفيين من ذوي الخبرة لإجراء أبحاث عن العلامات التجارية الإخبارية عبر الإنترنت من أجل توفير تقييمات للمواقع. قد يشكّل غياب أي تقييم رسمي ومستقل للمواقع أحد العوائق الرئيسية لمثل هذه المبادرات، خاصةً إذا كانت الجودة مشكوك فيها والمستخدمون غير مُدركين لذلك (Wilner, 2018 Funke & Mantzarlis, 2018b)

7.3.7 التحديات والفرص

تفضي هذه الأنواع من المبادرات، لا سيّما تلك القائمة على الاعتماد، إلى عدد من التحديات. ويرد نقاش أكثر تفصيلاً عن القضايا المحتملة التي تتمحور حول الأصالة وتصنيف المصداقية في تقرير Witness (Witness Media Lab, 2019)، غير أننا سنورد مُلخّصاً موجزاً أدناه. وتجدر الإشارة أولاً إلى أنّ التأثير الإجمالي الحالي لهذه المبادرات منخفض للغاية لأنها لا تُطبّق على نطاق واسع بعد، كما يتضح من عدد المرات القليل نسبياً التي جرى فيها تنزيل ملحقات كثيرة. فحتى NewsGuard الذي يضمّ 78,000 مستخدم في جميع أنحاء العالم لا يشكّل سوى نقطة في المحيط، سيّما وأنه يستهدف عامة الجمهور.

تقضي التحديات الأهم بالتصدّي لمسألة حياد الأطراف التي تتّخذ القرار بشأن المصداقية والثقة وكيفية قيامها بذلك، إلى جانب معالجة مسألتي التنوُّع والإقصاء. وتبرز أيضاً تحديات في مجال رصد الاعتماد وتجديده/إلغائه على مدى فترة زمنية. علاوةً على ذلك، وبما أنّ أدوات الاعتماد والأدوات

القائمة على الثقة تُنفّذ على نطاق أوسع، فقد أصبحت في حدّ ذاتها بيانات الثقة الفعلية عبر بيئات إعلامية متنوعة، مما يؤدي إلى استبعاد غير مقصود لبعض الكيانات الإعلامية وإدراج لبعضها الآخر. وإذا أصبحت المصادقة القاعدة التلقائية لوسائل الإعلام عبر الإنترنت، فإنّ ذلك يؤثر على التعددية في إطار حُريّة التّعبير، لا سيما بالنسبة للأشخاص المحرومين بالفعل في هذا المجال. وسيطرح هذا الأمر مشكلة بالنسبة إلى الأشخاص غير القادرين على الوصول إلى تكنولوجيا التدقيق أو الذين قد لا يرغبون في نشر معلومات يُحتمل أن تكون حساسة، كموقعهم مثلاً[484]. وهذا يشمل الأشخاص الذين هم في بلدان الجنوب، والذي يستخدمون هواتف تم تعطيل ميزة الحماية فيها[485]، والذين هم أيضاً في الغالب من النساء ويعيشون في المناطق الريفية أو النائية. وهناك أيضاً مشكلات محتملة تتعلق باستخدام الأصالة والتدابير القائمة على المصدر كسلاح متى كان استخدمها إلزامياً، لا سيما في سياق التشريعات المتعلقة «بالأخبار الزائفة».

علاوةً على ذلك، إذا قامت الشركات بتصنيف المصداقية، يبرز خطر أن تعمد منظمة تجارية ما إلى تحديد ما هو حزبي وما هو غير ذلك، الأمر الذي ينطوي على خطر اتخاذ قرارات غير شفافة ودمج تحيزات استراتيجية. كما يبرز خطر آخر، وهو أن تصبح هذه المقاربة «صيغة تُطبّق على الجميع»، من دون مراعاة الخصائص الثقافية والمجتمعية لبلدان معينة، وأن تدل ضمناً على أن بعض السرديات القائمة على الوقائع هي في جوهرها أكثر قيمةً من سرديات أُخرى تقوم على وقائع -بدلاً من الإشارة إلى مُحتوى غير قائم على الوقائع. وكثيراً ما يصعُب الحفاظ على الحياد في مثل هذه المبادرات، وقد يطرح تحديد ذلك على أساس مستمر، إشكالية معينة، مما يُثير أيضاً تساؤلات بشأن التواتر الدوري للمراجعة المستمرة لتصنيف المصداقية وآلياته. إلى جانب ذلك، تستطيع مبادرات تصنيف المصداقية أن تؤثر سلباً على التنوُّع المشروع في تفسير السرديات القائمة على الوقائع.

ومن المسائل ذات الصلة، دقة الأدوات المُستعملة لتصنيف المصداقية وشفافيتها بشكل عام. فعلم المصدر والتحليل الجنائي للمواد الإعلامية ذات الصلة ليسا بالأمر البسيط ولا يمكن تفسيرهما بسهولة. فحتى ولو بدا تصنيف مصادر المعلومات مع المعلومات عن المصدر حلاً بسيطاً، فإنّه ليس فحسب صعباً ومُعرّضاً للخطأ، بل لا يتضح دائماً متى يكون مغلوطاً. وقد تكون مسألة تصنيف المصداقية غير موضوعية، كما سبقت مناقشته، وقد تكون أيضاً عرضةً للخطأ سواء نُفذت يدوياً أو آلياً. وقد نوقشت قضايا الدقة المحيطة بأداة MediaBiasFactCheck في وقت سابق في هذا الفصل، وهي بالتأكيد ليست الأداة الوحيدة التي تُطرح نوعية نتائجها للنقاش. فمَن بالتالي ينبغي أن يُشرف على جودة هكذا أدوات؟

بعيداً عن موضوع الموثوقية، ماذا عن الاستجابات النفسية لمختلف الآليات، لا سيما تلك المتعلّقة بالتمكين. فعندما يوسم المُحتوى أو يُصنّف بطريقة ما، يبرز عدد من المخاطر المتعلقة بالتفسير. أولاً، إذا تم وسم عدد كبير من الإيجابيات الكاذبة (أي إذا اعتُبرت معلومات مشروعة أو وقائعية كثيرة مشبوهة أو غير جديرة بالثقة)، يعتاد الناس على ذلك ويميلون إلى الإفراط في تفسير الإنذارات الكاذبة، معتقدين أن الأدوات مفرطة الحساسية وأن تصنيف المصداقية ليس قابلاً للتصديق. من ناحية أخرى، إن لم يُوسَم المُحتوى بمصدر أو لم تُصنّف مصداقيته (على سبيل المثال، إذا لم يتّضح كيف ينبغي

---

[484] https://www.axios.com/deepfake-authentication-privacy-5fa05902-41eb-40a7-8850-5450bcad0475.html

[485] إن الهاتف الذي تم تعطيل ميزة الحماية فيه هو هاتف جرت قرصنته لتحريره من القيود التي فرضها المُصنّع. وبالتالي، لهذا الأمر تبعات على البرمجية، وخيارات التتبّع.. إلخ، والتي يمكن أن تتوفر فيه.

تصنيفه، وإذا كان لابدّ من تجنب الإيجابيات الكاذبة)، فقد يُفترض أن المُحتوى جدير بالثقة، الأمر الذي قد ينطوي على مخاطر (انظر: Pennycook et al.,2020). وبالنسبة إلى تزويد المستخدمين بالمعلومات، كالتفسيرات بشأن تصنيف المصداقية، أو حتى مجرد تفسيرات تساعدهم على اتخاذ قرار؛ لابدّ من الاختيار بين تقديم معلومات كافية ومفصّلة بما فيه الكفاية لتكون واضحة من جهة، وإدخال الكثير من التعقيدات، الأمر الذي يُديم الفجوة بين الخبير والمستخدم.

وبشكل أكثر عمومية، ما زالت مبادرات كثيرة من هذا النوع فتيّة للغاية، ولم يتم بعد اعتماد أي من مقاربات تصنيف المصداقية أو مقاربات المصدر أو «التقاط الصور الخاضع للسيطرة» controlled capture على نطاق واسع. من ناحيةٍ أخرى، وكما نوقش أعلاه، قد يفضي الاستخدام واسع النطاق لهذه الأدوات إلى مشاكل التحيز الاستراتيجي والاستبعاد والإدراك النفسي غير المقصود. وتُثار أيضاً تساؤلات أساسية حول كيفية نشر هذه المقاربات على نطاق أوسع: فهل سيتمّ ذلك بالتعاون مع المنصات أو في نظام بديل. والأمر ذو صلة بشكل خاص بالتكنولوجيات، كاستخدام قواعد البيانات المتسلسلة blockchain، علماً أن تطبيق هذه المقاربات خارج وسائل الإعلام والمؤسسات في بلدان الشمال محدوداً.

بالنسبة إلى أنظمة المصادقة على المُحتوى واسعة النطاق، تبرز بعض المسائل لفائدة إدارة تحديات اعتمادها على ضوء التبعات التقنية والمجتمعية. فلتتبّع المصدر، تُطرح أسئلة حول أسباب الخصوصية والتكتّم عن الهُويّة المشروعة. فنتساءل مثلاً: لماذا يختار الناس عدم الانخراط، أو يختارون الانخراط فقط في عناصر مختارة من المُحتوى، فضلاً عن القيود التقنية. وهذا يقود إلى سؤال حول كيفية التأكُّد من أن الثقة تقوم على أساس اختياري وليس إلزامي، وأنها بالتالي مجرد إشارة عن الثقة، وليست تأكيداً لها. وتشكّل النقطة الأخيرة معضلة هامة بالنسبة إلى الكثير من أنواع المبادرات المضادة للتضليل الأخرى- فلكي تتحلّى هذه الآليات بالفعالية، لابدّ لها أن تكون واسعة الانتشار، ولكن هذا الأمر يُسبّب مشكلات خطيرة عندما يختار الناس عدم المصادقة على بياناتهم، أو عندما تكون معلومات التدقيق ذات الصلة غير صحيحة أو مفقودة، لأسباب تتعلّق بالخصوصية المشروعة أحياناً، كالإبلاغ عن المخالفات.

وأخيراً، في ما يتعلق بتمكين المستخدمين، يُطرح عدد من الأسئلة حول أفضل الممارسات لإدارة المعلومات المعقّدة وعرضها. وكما نوقش أعلاه، لابدّ من تقديم المعلومات بطريقة بسيطة ولكن ذات مغزى حتى يتمكّن عامة الجمهور من استخدامها على النحو المناسب وفهم تبعاتها. غير أن العرض مفرط البساطة قد يؤدي إلى سوء التفسير؛ إذ يقترح أن مسألتي التدقيق والثقة مُتناقضتان. ومن ناحية أخرى، ولتفادي إرهاق المستخدم، يمكن أن تشكّل نظم الكشف التدريجي (عن طريق تقسيم المعلومات التفصيلية ذات الصلة المتعلقة بالثقة والمصداقية إلى مستويات أعمق يتعين استكشافها لفهمها بشكل إضافي) مقاربة مناسبة، ولكنها لم تُعتمد بعد. ومما لا شكّ فيه أنه ما من حل واحد يناسب الجميع لمشكلة تمكين المستخدمين من تمييز المعلومات التي يستهلكونها. ولا يزال يتعيّن إجراء المزيد من الأبحاث الاجتماعية والنفسية، فضلاً عن حلّ المسائل التقنية والقانونية.

وثمة تحدّ آخر يرتبط بالقضية النفسية، وهو المُفاضلة بين تمكين المستخدم من خلال تزويده بمؤشرات تدلّه على المعلومات المفيدة بشأن المصدر والثقة، وبالتالي تجنيبه التحيز الخارجي والمساعدة في الوقت نفسه على تثقيفه، وواقع أن العبء يقع الآن على المستخدم لاتخاذ القرارات، في حين قد لا

يكون مجهزاً بما فيه الكفاية لتفسير النتائج بشكل صحيح. وتنبغي الإشارة إلى أن تفسير التصنيفات يزيد من عبء المعالجة العصبية بالنسبة إلى المستهلك، علماً أنه عامل يفضي إلى انتشار المعلومات المُضلّلة والتحيز غير الواعي وغُرف الصدى filter bubbles. على العكس من ذلك، توفّر هذه الآليات أيضاً طريقاً إرشادياً مختصراً قد لا يكون دقيقاً (انظر مثلاً التقرير الحديث الصادر عن Witness الذي يناقش تاريخ علامات التحقُّق المُدقّق فيها، وكيف تُعتبر تفكيراً غرائزياً خاطئاً بدلاً من أن تكون تفكيراً عقلانياً (Witness Media Lab, 2019). أضف إلى ذلك أن الأدوات والممارسات التي تتيح للمستهلك التحقُّق من أنّ قسماً معيناً من المُحتوى جاء من مصدر محدّد لا تكون مفيدة إذا لم يكن المستهلك يفهم بشكل صحيح موثوقية ذلك المصدر. ويمكن القول بالتالي إنّ المقاربة الشاملة التي تتضمّن كلا الجانبين وتثقّف المستخدم ليتمكّن من التمييز في طريقة استخدامه للأخبار لا تزال بالغة الأهمية.

بالانتقال إلى الفرص، يمكن التصدي للكثير من التحديات المذكورة أعلاه من خلال الشفافية والتشاور واحترام التعددية والتنوع في إطار حُريّة التّعبير. علاوةً على ذلك، يتمثّل أحد أكبر نقاط القوة في استجابات التمكين، وفي استجابات تصنيف المصداقية بشكل خاص، في أنّ المؤشرات المنتجة سهلة التفسير ولا تحتاج إلى تدريب كثير. فأنظمة «إشارات المرور»، على سبيل المثال، توضح بشكل كبير ما هي الأمور الجديرة بالثقة وما هي الأمور المشبوهة. ويكتسي هذا الأمر أهمية خاصة بالنسبة إلى الجمهور الذي لا يمكن أن يُتوقع منه أن يصبح ملماً بالإعلام والمعلومات كالخبراء بين ليلةٍ وضُحاها، على الرغم من الفوائد التي توفّرها المبادرات التثقيفية، كما ذُكر في القسم السابق. ولا يمكن أن يُتوقع من الجمهور التحقُّق من الوقائع في جميع المحتويات التي يطّلع عليها. وبالتالي، تكمل الأدوات المساعِدة التي نوقشت أعلاه المهارات التي يوظّفها المستهلكون عند التعامل مع المُحتوى.

يمكن أن تؤدي هذه النظم أيضاً إلى فوائد طويلة الأجل، فيصبح مزودو الأخبار مثلاً أكثر جدارة بالثقة عموماً، لأنه متى تم تسليط الضوء على أوجه فشلهم مقارنةً بالجهات الفاعلة المعتمدة، يزداد الحافز للتحسّن. وفي الإطار نفسه، يُفضي اعتماد مجموعات من المعايير التي يمكن المصادقة عليها على نطاق واسع في صناعة الإعلام إلى فوائد محتملة، كالمساعدة في تعزيز الوضع الاقتصادي للناشرين الشرعيين.

بالإضافة إلى ذلك، تتيح مبادرات تتبّع المصدر التي تساعد المستخدم على فهم مصدر المادة وطبيعتها أو التي تسمح بالتدقيق في محتواها بطريقة ما توفير الوقت. وهذا أمر مهمّ للصحفيين في عالم الإعلام سريع الخُطى، ولكنه مهمّ أيضاً للجمهور العادي الذي لا يرغب في بذل الكثير من الجهد للتحقق من المصادر، حتى ولو أدرك أهمية ذلك. إلى جانب أدوات الثقة والمصداقية، تدفع هذه الميزة الموفّرة للوقت بدورها باتجاه اعتماد الممارسات الجيدة والمعايير من قبل شركات الإعلام الكبرى، ولكن أيضاً من قبل جميع منتجي المُحتوى الإعلامي، كالمدوّنين والمواطنين الصحفيين. وأخيراً، إذا حقّقت هذه الحلول النجاح المنشود، يمكن أن تعتمدها المنظمات الإعلامية على نطاق أوسع. ويمكن، على سبيل المثال، استخدام البروتوكولات القائمة على قواعد البيانات المتسلسلة blockchain لمشاركة البيانات الوصفية مع المُحتوى الإعلامي، أياً كان الاتجاه الذي يسلكه.

8.3.7 توصيات بشأن الاستجابات المتعلقة بالتمكين وبتصنيف المصداقية

يمكن بصورة عامة التشجيع على استخدام الوسائل المساعدة للمستخدم التي تنطوي على توحيد المعايير والمصادقة (دون المساس بالتعدُّديّة والتنوُّع)، فضلاً عن المقاربات التي يمكن تطبيقها على الصعيد العالمي وعبر مختلف المنصات.

وتؤدي التحديات والفرص المحددة أعلاه، وتبعاتها الهامة على حُريّة التّعبير، إلى التوصيات المحتملة التالية لاتخاذ إجراءات في هذه الفئة من الاستجابات.

توصيات لشركات ومنصات التواصل عبر الإنترنت ووسائل الإعلام الإخبارية:

- مع الاحترام الكامل للتعددية الإعلامية وتنوّعها؛ اعتماد معايير يمكن المصادقة عليها تتناول تصنيف مصداقية المؤسسات الإخبارية.
- النظر في مقاربات واضحة وبسيطة وموفّرة للوقت لتصنيف مصداقية المُحتوى، مع اتباع الشفافية الكاملة بشأن المعايير المعنية، وعملية التنفيذ، وفرص الطعن المستقلة.
- تجنّب الحلول السريعة التي قد تكون مضللة ولها عواقب غير مرغوب فيها، مثل دفع الناس إلى تصديق الوسم والمؤشرات التي قد لا تغطي كل جوانب القصة بشكل أعمى- أو تؤدي إلى تجاهل الناس لهذه الإشارات بسبب «الإيجابيات الكاذبة» أو التحيز.
- اختبار عملية وضع المعالم signposts والمؤشرات التي تشجّع الناس على التفكير في أنفسهم، وترفع مستوى الدراية الإعلامية والمعلوماتية لديهم.
- التأكّد من أن استجابات التمكين وتصنيف المصداقية تعمل جنباً إلى جنب مع الاستجابات التثقيفية لتحقيق أفضل النتائج.
- تنفيذ آليات أفضل لضمان الشفافية والمُساءلة في المؤسسات والمجتمعات المحلية العاملة على تصميم مقاربات التمكين وتصنيف المصداقية وتنفيذها، فضلاً عن تقييمها المستقل.
- وضع استجابات المصداقية بعناية كبيرة، مع إيلاء الاعتبار للبلدان الأقل نمواً، وشركات الإعلام والتكنولوجيا الأصغر حجماً، والمجتمعات المحرومة التي قد تتأثّر سلباً بالحلول غير المرنة التي لا تُراعي أوجه انعدام المساواة وتعددية وسائل الإعلام وتنوعها.

توصيات للباحثين والمجتمع المدني:

- اختبار تنفيذ الحلول العالمية واعتمادها (مثل: بروتوكولات قواعد البيانات المتسلسلة) لتتبع المصدر وتجنّب المقاربات المُجزّأة.
- تتبّع الممارسات داخل الشركات الإعلامية وشركات ومنصات التواصل عبر الإنترنت ككل، بما في ذلك تقييم أهمية البيانات الوصفية للمُحتوى بغض النظر عن المكان الذي يؤول إليه ذلك المُحتوى.

تقدّم هذه الدراسة تصنيفاً أصلياً للاستجابات للتضليل، يتناول كامل نطاق الاستجابات على الصعيد العالمي ويتوقّف عند عدد كبير من المبادرات والجهات الفاعلة. علاوةً على ذلك، يقدم البحث مقاربة فريدة من نوعها؛ إذ يُدرج التحديات والفرص المتعلقة بحُريّة التّعبير في صميم التحليل.

والجديد في هذه الدراسة أنها تركّز على تحديد 11 نوعاً مختلفاً من الاستجابات للتضليل وشرحها، وهي استجابات تم تقييمها على ضوء الموضوعات التي تركّز عليها، بدلاً من النظر فيها استناداً إلى الجهات الفاعلة الرئيسية المعنية. وتماماً كذلك، يكتسي هذا المشروع نطاقاً عالمياً؛ إذ ضمّت الدراسة عدداً من المبادرات التي اعتمدتها البلدان النامية حرصاً على ضمان التنوع الجغرافي.

بالإضافة إلى ذلك، أتاحت الجنسيات والتخصُّصات المتنوعة للباحثين العاملين على هذا المشروع بروز وجهات نظر متعددة وتلاقيها، مما أدى إلى إنتاج بحث سياساتي غني ومهم يرتبط بكل من الممارسة والتأثير، ويركّز على التدابير التكنولوجية، وتدخّلات الدول، والمبادرات التربوية، والتدخّلات الحكومية والصحفية.

وأخيراً، تسعى الدراسة إلى تفكيك المعلومات المُضلّلة بطريقة مُبتكَرة؛ من خلال التحقيق في أُسُس هذه الاستجابات، لا سيما نظريات التغيير الكامنة وراءها، فضلاً عن تحليل أهدافها ومصادر التمويل التي تعتمد عليها.

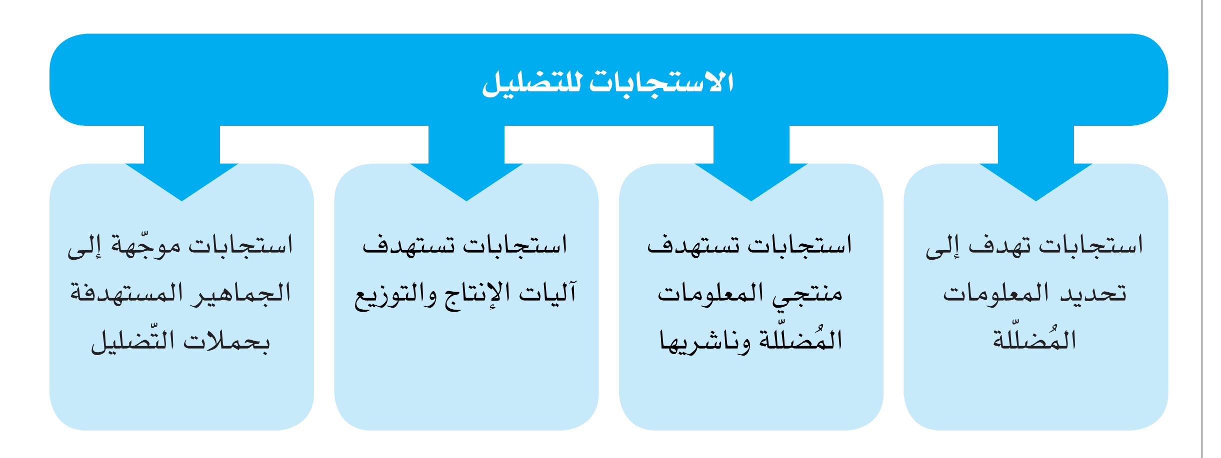


الفصل الثامن: التحديات والإجراءات المُوصَى بها

المؤلفتان: جولي بوزيتي وكالينا بونتشيفا

1.8 تلخيص التصنيف

يبدو من المناسب، في هذه المرحلة، تلخيص التصنيف الأصلي للاستجابات التي وضعها الباحثون لهذا التقرير. وقد تم أولاً جمع أنواع الاستجابات ضمن **أربع فئات شاملة**:



وجرى بعد ذلك **تحديد 11** طريقة منفصلة للاستجابة في إطار هذه الفئات الشاملة الأربع:

1. **استجابات تهدف إلى تحديد المعلومات المُضلّلة** (هدفها تحديد المعلومات المُضلّلة وكشف زيفها وإظهار حقيقتها)
   - الرصد والتحقُّق من الوقائع
   - استجابات استقصائية

2. **استجابات تستهدف منتجي المعلومات المُضلّلة وناشريها من خلال تغيير البيئة التي تحكم سلوكهم وتشكّله** (استجابات قانونية وسياسية)
   - استجابات تشريعية وما قبل تشريعية وسياساتية
   - الحملات الوطنية والدولية المناهضة للمعلومات المُضلّلة
   - الاستجابات الانتخابية

3. **استجابات تستهدف آليات الإنتاج والتوزيع** (تتعلّق بسياسات وممارسات المؤسسات التي تتوسّط في المُحتوى)
   - استجابات متعلّقة بتنظيم المُحتوى
   - استجابات تقنية وخوارزمية
   - استجابات اقتصادية

4. **استجابات موجّهة إلى الجماهير المستهدفة بحملات التّضليل** (موجّهة إلى «الضحايا» المحتملين للمعلومات المُضلّلة)

- استجابات معيارية وأخلاقية
- استجابات تثقيفية
- استجابات التمكين وتصنيف المصداقية

تتّسم هذه الفئات المختلفة من الاستجابات بطبيعة تآزُرية وتكافلية. فهي تعمل بشكل منفصل ولكن أيضاً مترابط لمواجهة المعلومات المُضلّلة. وتدعم الاستجابات المعيارية والأخلاقية، على سبيل المثال، الكثير من أنواع الاستجابات الأخرى، في حين تؤدي استجابات الرصد والتحقُّق من الوقائع والاستجابات الاستقصائية دوراً أساسياً في توجيه الاستجابات الاقتصادية والتمكينية واستجابات تنظيم المُحتوى.

2.8 نظرة شاملة مواضيعية

يبيّن مسح خارطة المشهد وتحليل الثغرات البحثية في الفصل الثالث وفرة الدراسات المتعلقة بأساليب التّضليل والتصدي للمعلومات المُضلّلة. غير أنّ الأبحاث التي أُجريت قبل تاريخ إعداد هذا التقرير تُركّز إلى حد كبير على سياقات الدول الغربية والدول المتقدمة الناطقة باللُّغة الإنكليزية.

وبصورة عامة، في حين تتوفّر مجموعة متنامية من الأبحاث والبرمجيات والتدريبات والموارد المعرفية التي تركّز على معالجة المعلومات المُضلّلة، يُسجّل غياب مقارن لتلك التي تركز على التّضليل من منظور حقوق الإنسان، وحُريّة التّعبير، وتزايد فرص الوصول إلى تكنولوجيا النطاق الواسع واستخدامها في جميع أنحاء العالم.

وفي الوقت نفسه، تُسجّل ندرة في الأبحاث المتعلقة بآثار التعرّض للتضليل والمُحتوى المناهض للتضليل على الأشخاص المعرّضين لهما. ويحدّد تحليل الثغرات البحثية أيضاً هُوّة تفصل الأبحاث الأكاديمية والتحقيقات الصحفية عن تلك التي يطلب المجتمع المدني والمنظمات الحكومية الدولية إجراءها. بالإضافة إلى ذلك، يُلاحظ أن التعاون بين هذه القطاعات نادر، على الرغم من القيمة الكبيرة التي قد ينطوي عليها. ويبدأ هذا التقرير بملء الفراغ المتعلّق بالأبحاث التي تتناول الاستجابات للتضليل في العالم النامي، وهو يُعدّ نقطة انطلاق للمساعي البحثية المستقبلية التي تُركّز على التنوُّع اللُّغوي والجغرافي وتنوّع التخصصات.

من خلال تحليل الاستجابات الرامية إلى تحديد المعلومات المُضلّلة (الرصد، والتحقُّق من الوقائع، والاستجابات الاستقصائية) من حول العالم (القسمان 4.1 و4.2)، تسلّط الدراسة الضوء بوجه خاص على قيمة التعاون عبر الحدود والتعاون متعدد التخصصات وبين أصحاب مصلحة متعددين في المبادرات الرامية إلى تحديد الأكاذيب والتلفيق.

ومما لا شكّ فيه أن الأحداث الكبرى، مثل الانتخابات وحالات الطوارئ الصحية العامة، تشكّل فرصة لتطوير استجابات مستقلة تتيح تحديد المعلومات المُضلّلة من أجل إعادة التأكيد على قيمة الوقائع، وتشجيع التفكير العام بالمُحتوى الذي يعتبره الناس موثوقاً وما يُقررون مشاركته. وتكتسي الاستجابات الآيلة إلى تحديد المعلومات المُضلّلة أهمية أيضاً في رصد تقاطع المعلومات المُضلّلة مع خطاب الكراهية– ضد النساء والأقليات والمهاجرين وغيرهم من المواطنين والمجتمعات المُستضعفة، بما في ذلك الأطفال والمسنّين الذين قد يكونون عُرضةً للتضليل بشكل خاص.

وكما تمّ التطرّق إليه، يطرح الحجم الهائل من المعلومات المُضلّلة، ومجموعة مصادرها، وأنواعها المختلفة صعوبة في مراقبة المشكلة وكشفها وتقييمها وزيادة الوعي بها. وتُعدّ التحديات كثيرة ومعقّدة– ومن أهمها ضرورة العمل على نطاق واسع، وبلغات متعددة وعبر عدد لا يُحصى من البلدان في الوقت الحقيقي. وتُمثّل أيضاً الحاجة الموازية إلى تحقيق الأثر عن طريق الحدّ من المعلومات المُضلّلة ووقفها بطرق قابلة للقياس تحدياً رئيسياً، شأنها شأن صعوبة تقييم هذه الجهود.

يكمن أحد التحديات الخاصة التي لابدّ من التوقف عندها في دور الدول والجهات الفاعلة السياسية بوصفها ناقلة للمعلومات المُضلّلة ومُستجيبة لها على حد سواء. غير أنّ هذه الدراسة تركّز على وصف الاستجابات للمعلومات المُضلّلة وتقييمها، وبالتالي لا تشدّد على دور هذه الجهات الفاعلة كمصادر للمعلومات المُضلّلة وجهات عاملة على تضخيمها. ومع ذلك، لابدّ من الإشارة إلى أنّ انتشار هذه المصادر كمُحرّضة على التّضليل ومُروّجين له يؤكد على الواجب الملقى على عاتقها والمتمثّل بالاستجابة للأزمة بطرق شفافة وخاضعة للمساءلة.

ويتم أيضاً في هذه الدراسة التأكيد على أهمية الدفاع عن حقوق حُريّة التّعبير جنباً إلى جنب مع الاستجابة للمعلومات المُضلّلة، من خلال الحرص على عدم عرقلة جهود التحقُّق من الوقائع بفعل شراكات أقامتها شركات التواصل عبر الإنترنت التي قد تسعى إلى الحدّ من نطاق التحقُّق من الوقائع أو استخدام مبادرات من هذا القبيل كتغطية في مجال العلاقات العامة، ولا من جانب الدولة والجهات الفاعلة السياسية التي تتصرّف كمصادر أولية للمعلومات المُضلّلة وجهات رئيسية تعمل على تضخيمها.

ومن المسائل الأخرى التي تمّ النظر فيها فشل جهود التدقيق والتحقُّق من الوقائع في أوساط بعض ناشري الأخبار –إما كنتاج للحزبية المفرطة، أو سيطرة الدولة، أو «استيلاء المنصات»، أو ضعف المعايير أو النقص في الموارد– الأمر الذي يمكن أن يحوّلهم إلى ناقلين للمعلومات المُضلّلة ويؤدي إلى تآكُل ثقة الجمهور في العملية. لهذا السبب، يجب تطوير القدرة على التحقُّق من الوقائع والاضطلاع باستقصاءات –داخل قطاع الصحافة وفي أوساط منظمات المجتمع المدني التي يمكن أن تساعد في تعزيز جهود التنظيم الذاتي وتحسين المُساءلة المهنية.

ويمثّل هذا الأمر فرصةً لتعزيز الاستجابات الآيلة إلى تحديد المعلومات المُضلّلة على نطاق واسع. وفي حين تعهّدت بعض الشركات والمنظمات غير الحكومية بتقديم بعض التمويل للمنظمات المعنية بالتحقُّق من الوقائع، تبرز الحاجة إلى تأمين دعم مُستمر طوال فترات الانتخابات والأوبئة الحرجة وبعدها. فعلى سبيل المثال، من الواضح أن أهمية التحقُّق من الادعاءات المتعلقة بالهجرة والعِرق والتلقيح وتغيّر المناخ ستزداد مع الوقت.

بالنسبة إلى الاستجابات للتضليل التي تضطلع بها الدول- بما في ذلك الحملات ما قبل التشريعية والتشريعية والقائمة على السياسات العامة وحملات التثقيف العام والاستجابات الانتخابية (الأقسام 5.1 إلى 5.3) - فإن المبادرات تتقاطع. وتكون الجهات الفاعلة التي تُعتبر احتيالية أو مسيئة من بين الأهداف الرئيسية للاستجابات التنظيمية للحكومات، علماً أن هدفها المعلن يتمثّل في قمع التحريض واسع الانتشار على الكراهية والعنف، والحماية من المخاطر التي تحدق بالأمن القومي، ومنع تعطيل الانتخابات الديمقراطية، وتجنّب التداعيات الجيوسياسية. وقد تشمل هذه الاستجابات أيضاً الاستثمار في مبادرات التحقُّق من الوقائع، والصحافة ذات المصلحة العامة، ومشروعات الدراية الإعلامية والمعلوماتية. بالإضافة إلى ذلك، تركّز هذه الاستجابات على شركات ومنصات التواصل عبر الإنترنت، وتستهدف قوتها الاقتصادية وسلوكها التقني. واستناداً إلى افتراض مفاده أنّ البنية (التوزيع بين الأقران) والخوارزميات المُستخدمة تُتيح تضخيم المعلومات المُضلّلة، تسعى مبادرات تنظيمية كثيرة إلى التشجيع على تحمّل هذه الجهات الفاعلة مسؤولية متزايدة.

وقد تأتي التدخّلات على شكل دعم للتنظيم الذاتي عبر إجراءات تنظيمية قاسية يمكن أن تؤدي إلى أوامر بإنهاء الخدمة وحجب المُحتوى. وقد يتم استهداف المروّجين المحترفين للتضليل، مثل وكالات العلاقات العامة المتخصّصة في المعلومات المُضلّلة واسعة الانتشار، تماماً كالسياسيين أنفسهم من خلال موجبات جديدة تتعلق بالشفافية في الإعلان السياسي عبر الإنترنت ودعم مبادرات التحقُّق من الوقائع خلال الانتخابات.

إنّ الوتيرة الهائلة للتغيّر التكنولوجي تجعل محاولات عديدة للتنظيم بالية قبل تطبيقها حتى. ومن ناحية أخرى، لابدّ من الإشارة إلى أنّ الرغبة في التحرك بسرعة للحد من المعلومات المُضلّلة واسعة الانتشار دون إجراء أي مناقشة وتدقيق وشفافية مناسبة؛ تنطوي على مخاطر كبيرة تحدق بحقوق الإنسان، لا سيما حُريّة التّعبير وحُريّة الصحافة والنفاذ إلى المعلومات واحترام الخصوصية. ويمكن أن يؤدي هذا الأمر إلى عواقب غير مرغوب فيها، مثل تجريم الصحافة الفعليّ من خلال ما يُسمّى بقوانين «الأخبار الزائفة» وغيرها من التدابير التي قد تصيب عن غير قصد الصحافة المشروعة.

تتناول الأقسام من 6.1 إلى 6.3 الاستجابات الرامية إلى الحد من إنتاج المعلومات المُضلّلة وتوزيعها، وهي الاستجابات التنظيمية، والاستجابات التقنية والخوارزمية، والاستجابات الاقتصادية. ويمكن القول إنّ التحدي الأكبر الذي تم تحديده في مجال هذه الاستجابات هو أنه على الرغم من الاعتراف بالدور الذي لابدّ أن تؤديه شركات التواصل عبر الإنترنت للحدّ من المعلومات المُضلّلة المنشورة على منصاتها، تنجم مشكلات محتملة عن تفويض السلطة التنظيمية (الذاتية) إلى هذه الشركات الخاصة. هذا ما يحصل بالتحديد عندما يفضي تفويض كهذا إلى تقليص المساءلة والطابع القضائي للقرارات ذات الصلة بالتّعبير التي تكون عادةً من مسؤوليات الدول، والتي ينبغي أن تمتثل للقوانين والمعايير الدولية لحقوق الإنسان. ويكمن الخطر في إمكانية خصخصة الرقابة بصورة فعلية.

في السياقات الديمقراطية، قد تنصّ اللوائح التنظيمية صراحةً على مثل هذا التفويض. وفي هذه الحالة، يمكن أن تتوفّر مُساءلة عامة. ومع ذلك، تُطلق يد الشركات إلى حد كبير لتقوم بتنظيم المُحتوى ذاتياً، لمجموعة من الأسباب السياسية والاقتصادية والتكنولوجية. وتعكس هذه المسألة الحاجة المُلحّة إلى اعتماد عملية طعن قوية وإلى توحيد المعايير التي تتبعها عملية وضع تقارير الشفافية التي تتناول طريقة

اتخاذ مثل هذه القرارات– الآلية إلى إزالة المُحتوى وتركه على حد سواء في حال الإساءة المستهدفة والتّضليل، على سبيل المثال. وتزداد هذه المشكلة حدةً لأن شركات ومنصات التواصل عبر الإنترنت تلجأ بشكل متزايد إلى الميكنة والخوارزميات القائمة على الذكاء الاصطناعي بصفتها وسيلة فعالة من حيث التكلفة للسيطرة على المعلومات المُضلّلة على نطاق واسع وبسرعة أقرب إلى الوقت الحقيقي. وبما أنّ الخوارزميات تخضع لكلٍّ من التحيز الضمني والصريح المحتمل في تصميمها وفي بيانات التدريب المُستخدمة لتطويرها (مع ما يترتب على ذلك من تبعات بالنسبة إلى المساواة بين الجنسَين والمساواة العِرقية)، فإنّ الأمر يطرح بشكل متزايد مشاكل كبيرة، لا سيما متى كانت الشركات قد حدّت أيضاً من قدرة المستخدمين على اللُّجوء إلى عملية طعن يديرها إنسان. ولكن من الناحية الإيجابية، تكون الاستجابات التكنولوجية، على مستوى مُحتوى محدد، أقل عُرضةً للضغوط الخارجية التي تُمارس (من قبل الدولة مثلاً) على المشغّلين البشر داخل الشركة لاتخاذ إجراءات معيّنة بشأن حالة تضليل مزعومة.

من جهتها، تُعتبر الاستجابات الهادفة إلى حجب الربح المالي والمرتبطة بالإعلانات نوعاً معيناً من الاستجابات التكنولوجية التي تركّز تحديداً على الحد من استحداث المعلومات المُضلّلة المنتجة لأغراض تحقيق الربح ونشرها. وهنا أيضاً، تماماً كما هو الحال مع الاستجابات التكنولوجية، تضطلع بغالبية الاستجابات الهادفة إلى حجب الربح المالي جهات فاعلة خاصة تتّخذ قرارات غير متّسقة ومُبهمة. وفي هذه الحالة، تتمثل المشكلات في عدم كفاية الشفافية التي توفرها شركات التواصل عبر الإنترنت في مجال الإعلانات، مما يحول في الواقع دون التدقيق والرقابة المستقلَّين. ولا تُلاحظ هذه المشكلة بشكل حاد عبر الكثير من المنصات والبلدان في مجال المعلومات المُضلّلة الصحية فحسب (مثل جائحة فيروس كورونا واللقاحات) أو الإعلانات المستندة إلى قضايا، بل أيضاً في مجال الإعلانات السياسية. وقد طُرحت هذه المشكلة بشكل خاص بالنسبة إلى فيس بوك، الذي رفض إلغاء الاستهداف الدقيق لمستخدميه في الإعلانات السياسية، وقاوم التدابير الرامية إلى إخضاع الخطاب السياسي المباشر للسياسيين (وأحزابهم والمنتسبين إليها) للتحقّق من الوقائع، خاصةً في الولايات المتحدة (Suarez, 2020).

وأخيراً، بحثت هذه الدراسة في الاستجابات الموجّهة إلى الجماهير المستهدفة بحملات التّضليل، أي الاستجابات المعيارية والأخلاقية، والاستجابات التثقيفية، واستجابات التمكين وتصنيف المصداقية (الأقسام 7.1 إلى 7.3). بالنسبة إلى الاستجابات في الفئة الأولى –أي المعيارية والأخلاقية– يمكن القول إن تطبيقها قد يكون بسيطاً ومقبول الكلفة نسبياً (وإن كان من الصعب تقييم تأثيرها)، وأنه بإمكانها أن تُعتبر سرديات مضادة تجذب البوصلات الأخلاقية للأفراد، أو تبدو متوافقة مع القيم الثقافية كمناهضة العنصرية أو مُكافحة كراهية النساء. ومع ذلك، تبرز مشكلة عندما لا ترتبط البوصلات الأخلاقية والمعايير المجتمعية بمبادئ الوصول إلى المعلومات وحُريّة التّعبير وحُريّة الصحافة والخصوصية– على النحو المنصوص عليه في القانون الدولي لحقوق الإنسان.

فعلى سبيل المثال، يبرز أحد أهم المخاطر المرتبطة باستجابات الدول للتضليل في المقاربات القانونية والتنظيمية التي تتعارض مع المعايير الدولية لحُريّة التّعبير (وحُريّة الصحافة المرتبطة بها) واحترام الخصوصية. وتتوفّر أيضاً واحدة من أكبر الفرص أمام شركات ومنصات التواصل عبر الإنترنت للارتقاء فوق المخاوف المتعلقة بالربح والنفوذ وإدارة السمعة من أجل اتخاذ إجراءات حقيقية تتيح التصدي للمعلومات المُضلّلة السياسية التي تهدد الديمقراطية، بالتوازي مع المعلومات المُضلّلة الصحية التي تُهدّد الحياة. ولكن على الرغم من ذلك، قد تُنزع صفة الشرعية، في نهاية المطاف، عن الاستجابات

المعيارية والأخلاقية المشروعة للتضليل من قبل الأفراد والمنظمات والدول الذين يختلفون حول النية من وراءها، بالطريقة نفسها التي يمكن تشويه الصحافة ذات المصداقية من خلال وصفها بـ«الأخبار الزائفة» من قبل أولئك الذين يسعون لتجنّب المُساءلة. لهذا السبب بالذات، يتعيّن على هذه الاستجابات أن تترافق وتدخلات إبداعية في مجال الدراية الإعلامية والمعلوماتية، وتدخلات مصمّمة لتمكين مستخدمي وسائل التواصل الاجتماعي ومتلقّي الأخبار.

أما بالنسبة إلى الاستجابات التثقيفية، فتستهدف التدخلات طائفة واسعة من الفئات العمرية- من تلاميذ المدارس إلى طلاب الجامعات، مروراً بالصحفيين والمواطنين المُسنّين، على الرغم من أن الوصول الشامل إلى هذه الفئات لم يتحقّق بعد في الكثير من البلدان. ويتمثّل أحد التحديات والفرص الأكبر في جعل التعليم من أجل المواطنة العالمية والدراية الإعلامية والمعلوماتية في مُتناول الأطفال في جميع أنحاء العالم. إنها لمهمّة تتطلّب استثماراً كبيراً في التدخلات (التي يمكن تطويرها بالطريقة الفُضلى بالتشاور مع أصحاب مصلحة مُتعددين، بما في ذلك شركات التواصل عبر الإنترنت)، وتثقيفاً مناسباً للمُعلّمين من خلال التدريب، فضلاً عن تثقيف أولئك الذين يعملون بالفعل في الميدان. وثمة تحدٍ هام آخر يتمثل في ضمان تصميم التدخلات الخاصة بالدراية الإعلامية والمعلوماتية والتعليم من أجل المواطنة العالمية بشكل كلي، والحرص على تضمينها قضايا مثل خطاب الكراهية والسلامة الرقمية والدراية الأمنية، فضلاً عن الهُويّة وحقوق الإنسان والاقتصاد السياسي لنماذج الأعمال الخاصة بالإعلام وشركات الإنترنت. وتواجه التدخلات التثقيفية أيضاً التحدي المتمثل في إدماج فهم دور الصحفيين والمدافعين عن حقوق الإنسان، لا سيما النساء والمجتمعات المهمشة، الذين يُستهدفون بشكل متكرّر في حملات التّضليل. وعلى نطاق أوسع، من المهمّ أيضاً الحرص على أن ينصبّ تركيز المجتمع على حلّ الأسباب الجذرية لمشكلة التّضليل، بدلاً من مجرد بناء القدرة على التكيّف معها كما لو كان ذلك يشكّل حلاً قائماً بذاته.

أخيراً، لابدّ من الإشارة إلى أنّ تحليل الاستجابات المتعلقة بالتمكين وتصنيف المصداقية قد كشف عن تحديات هامة ترتبط بضرورة معالجة التنوع والاستبعاد والحياد، مع الإشارة إلى الجهة التي تحدد ما هو جدير بالثقة ويتمتّع بالمصداقية وكيف تجري الإشارة إليه. وتبرز مخاوف خاصة تتمحور حول تبعات هذه النظم على تعددية وسائل الإعلام، على ضوء الطرق التي يمكن من خلالها شمل وسائل إعلام أو استبعادها (لا سيّما إذا ما كانت تعتمد على الوصول إلى تكنولوجيات معينة أو مجموعات معينة من المهارات)، وإمكانية استخدامها كسلاح في سياق قوانين «الأخبار الزائفة». وقد اعتُبر أيضاً أن رصد هذه النظم وتقييمها وتحديثها على المدى الطويل لمجال مثير أيضاً للقلق.

ما زالت استجابات كثيرة للتضليل ورد وصفها في هذه الدراسة حديثة نسبياً ولم تُعتمد بعد على نطاق واسع. ويرجع ذلك في بعض الحالات إلى أنّ بعض التكنولوجيات لا تزال قيد التطوير أو التكييف، لأنها لم تجد قوة دفع واسعة النطاق بعد، أو لأن عوائق قانونية أو أخلاقية اعترضت سبيلها. فعندما لا تُستخدم مثلاً مقاربات تصنيف المصداقية على نطاق واسع، تتأثّر فعاليتها بوضوح ويكون فهم إمكانياتها محدوداً. والمسألة هذه توضح «معضلة كولينجريدج» (Collingridge, 1980) التي تفترض أساساً أنه لا يمكن بالإجمال التنبؤ بالعواقب الاجتماعية للتكنولوجيا إلى حين تطوير التكنولوجيا بالفعل، وعندها غالباً ما يكون قد فات الأوان، أو بات التغيير أكثر صعوبةً على الأقلّ. وفي هذه الدراسة، سُجّلت محاولة للردّ على «معضلة كولينجريدج» على مستوى وضع السياسات من خلال التقارب بين الخبرات متعددة التخصصات والتجارب العمليّة التي تركّز على تقييم الاستجابات للتضليل استناداً إلى حقوق الإنسان.

3.8 تطبيق التصنيف على جائحة التّضليل

تـم اسـتخدام تصنيـف الاسـتجابات للمعلومـات المُضلّلـة الـوارد بالتفصيـل فـي هـذا التقريـر البحثـي فـي موجزَيـن للسياسـات نشـرتهما اليونسـكو حـول الاسـتجابات للمعلومـات المُضلّلـة المتعلّقـة بجائحـة فيـروس كورونـا بغـرض تقييـم مـدى انطبـاق الصيغـة علـى أزمـة تضليـل محـدّدة (Posetti & Bontcheva, 2020a; 2020b). وفـي سـياق أزمـة المعلومـات المُضلّلـة حـول جائحـة فيـروس كورونـا، التـي وصفهـا موجـزا السياسـات بأنهـا «مشـكلة تضليـل»، تـم تحليـل القضايـا المشـتركة والمتقاطعـة، وتـم تشـريح الافتراضـات والتحديـات والفـرص الكامنـة وراء كل نـوع مـن أنـواع الاسـتجابات بشـكل منهجـي، ممـا يوفـر إطـاراً تعليميـاً لهـذه الدراسـة التـي تتّسـم بطابـع أكثـر عموميـةً وشـمولاً.

وقـد أظهـرت هـذه العمليـة إمكانيـة تطبيـق الإطـار علـى مشـكلات تضليـل أخـرى محـددة، كالمشـكلات المرتبطـة بالانتخابـات وحمـلات مُكافحـة اللقـاح وتغيّـر المنـاخ.

4.8 التقييم الشامل لعدة قطاعات

تسـتند الاسـتجابات التـي تـم تقييمهـا فـي هـذه الدراسـة إلـى افتراضـات أساسـية، بعضهـا موضـع تسـاؤل ويسـتدعي التدقيـق. وقـد تكـون هـذه الافتراضـات ضمنيـة بـدلاً مـن أن تكـون صريحـة فـي بعـض الحـالات، وقـد تـؤدي فـي حـالات أخـرى إلـى تقويـض النتائـج المرجـوة مـن التدخـلات. إلـى جانـب ذلـك، قـد تتغاضـى بعـض الافتراضـات عـن قضايـا حقـوق الإنسـان، فـي حيـن قـد يفتـرض البعـض الآخـر، مـن بـاب الخطـأ، أنهـا لـن تُفضـي إلـى آثـار غيـر مقصـودة تمـسّ هـذه الحقـوق.

يسـعى الكثيـر مـن طـرق الاسـتجابة المعروضـة هنـا إلـى تعزيـز وضـوح المعلومـات الحقيقيـة ذات المصلحـة العامـة وتقويتـه (مثـل: الصحافـة المسـتقلة، والمعلومـات المشـروعة عـن الصحـة العامـة، أو جهـود حمايـة الانتخابـات). وتهـدف طـرق أخـرى إلـى القضـاء علـى المعلومـات المُضلّلـة (أو علـى الأقل التقليـل مـن بروزهـا أو طابعهـا الرسـمي الزائـف). وهنـاك أيضـاً أمثلـة عـن اسـتجابات تهـدف إلـى ممارسـة الرقابـة السياسـية أو مقاومـة التنظيـم.

وفـي حيـن يمكـن تقييـم طبيعـة مشـكلة التّضليـل وآثارهـا بشـكل مختلـف حـول العالـم ومـن قِبـل جهـات فاعلـة مختلفـة، فـإن جميـع التدخـلات المقدّمـة هنـا قـد صُمّمـت لإحـداث تغييـر[486]. ولهـذا السـبب، تتضمّـن كلّهـا «نظريـة التغييـر»، علمـاً أن الأمـور التـي تسـعى إلـى تغييرهـا والسـبب وراء ذلـك وأسـباب التصـرّف قـد تتنـوّع وتتعـدّد. فعلـى سـبيل المثـال، فـي حيـن تتخـوّف الـوكالات الإخباريـة مـن تأثيـر التّضليـل علـى قيمـة صحافتهـا (بسـبب الهجمـات عبـر الإنترنـت المصمّمـة لتشـويه سـمعة التقاريـر النقديـة، مثـلاً) وآثـار تـآكل الثقـة المترتبـة

486 ملاحظة: نُشرت أفكار وجمل كثيرة ترد في ما يلي في موجزَي السياسات اللذين أعدّتهما اليونسكو بشأن «مُشكلة التّضليل»: https://en.unesco.org/covid19/disinfodemic

على نموذج الأعمال، فإنّ شركات التواصل عبر الإنترنت لا ترى بالضرورة التّضليل كمشكلة مرتبطة بالاقتصاد؛ وإنما مسألة تتعلّق بالعلاقات العامة ومشكلة تنظيمية محتملة. وقد ترغب الحكومات في فرض لوائح تنظيمية لأسباب مختلفة، من بينها أنها غير راضية عن استجابات الشركات، أو لأنها ترى فرصة لتهدئة الصحافة المستقلة النقدية من خلال جهود مُكافحة التّضليل المزعومة، كقوانين «الأخبار الزائفة». وعلى الرغم من أنّ «نظرية التغيير» التي تقف وراء هذه التدخلات عادةً ما تكون غير مفصّلة، فإنّ نقاط القوة والضعف في النظرية التي يتم الاعتماد عليها تُعدّ أساسية لضمان فعالية التدخلات، وللتصدي لأي آثار غير مقصودة.

إنّ تطبيق التصنيف الذي تم تطويره في إطار هذا التقرير البحثي على مُشكلات معلومات مضللة محددة، مثل دراسة الحالة حول «مشكلة التّضليل» بشأن جائحة فيروس كورونا المُشار إليها في مختلف أقسام هذه الدراسة، كشف عن أوجه القصور في مجرد إعادة استخدام الاستجابات المضادة للتضليل القائمة (كتلك المصمّمة لحماية الانتخابات، على سبيل المثال) مع أنواع ومظاهر جديدة للتضليل. وفي حال الاستجابات لمشكلة التّضليل المرتبطة بجائحة فيروس كورونا، افتُرض أن نماذج التدخّل الموجودة من قبل (كتلك المطبّقة على إنكار تغير المناخ وحركة مُكافحة اللقاح) تشكّل قاعدة كافية للاستجابات التي تتصدّى للمعلومات المُضلّلة كثيفة الانتشار والمميتة المرتبطة بالوباء. ولكن في كثير من الحالات، لم تكن هذه النماذج القائمة مناسبة للغرض بسبب النطاق العالمي للتأثيرات المرتبطة بالجائحة التي ولّدت مثل هذا الارتباك الهائل وعدم اليقين وبسبب سرعتها ومداها. وكان لابدّ من اعتماد تدخلات أكثر تنسيقاً وتكاملية، عبر طائفة أوسع من الجهات الفاعلة، لمواجهة «مشكلة التّضليل».

يكمن افتراض أساسي وراء الكثير من المبادرات التي نُفّذت استجابةً للتضليل، وهو يفيد بأن المبادرات تستند إلى الحدس بشأن ما هو مطلوب، وكيف يُتوقّع أن ينجح التدخل. والسبب هو أن المبادرات تُطبّق في غياب أدلة تجريبية. ويتّضح أن عدداً قليلاً من الجهات الفاعلة التي تتعامل مباشرةً مع المعلومات المُضلّلة تلحظ الإشراف المستقل أو تقييم الأثر على المدى الطويل، بما في ذلك رصد الآثار غير المقصودة وتقييمها. ومن أهم المخاطر الناشئة عن هكذا تصرّف تقويض الحق في حُريّة التّعبير على المدى الطويل، بما في ذلك حُريّة الصحافة، والنفاذ إلى المعلومات، وحماية واحترام الخصوصية.

غير أن هناك مسألة أخرى تتمثل في أن المُساءلة عن بعض الاستجابات ليست دائماً واضحة أو شفافة. ويتّضح أيضاً أن الكثير من الاستجابات لا تراعي المعايير الدولية ذات الصلة بالقيود المفروضة على حقوق حُريّة التّعبير، لا سيما تلك الخاصة بالضرورة والتناسب. ويؤدي هذا التجاوز إلى انتهاك الحق المشروع في حُريّة التّعبير، لا سيما حُريّة الصحافة التي تشكّل شرطاً مسبقاً لتوفير المعلومات القادرة على المساعدة في التغلب على تحدي التّضليل.

وقد تفاقمت قضايا المساءلة هذه في سياق جائحة فيروس كورونا التي تطلّبت استجابات تُصمّم وتنفّذ بسرعة في ظل ظروف طارئة، من أجل التعامل مع التهديد العالمي غير المسبوق الذي يحدق بالصحة العامة ويترافق وتشعبات اجتماعية واقتصادية هائلة تضخّمها مشكلة التّضليل.

5.8 استعراض التحديات والفُرص

- الأُطُر الزمنية: تُوجَّه بعض الاستجابات –مثل اللوائح التنظيمية الجديدة– نحو تحقيق نتائج فورية، بينما تكون استجابات أخرى، كتمكين المستخدمين، متوسطة الأجل. بعد ذلك، هناك تدابير مثل تطوير الدراية الإعلامية والمعلوماتية التي يستغرق ترسيخها وقتاً أطول، غير أنها قد تفضي إلى نتائج دائمة. وتتوفّر تدابير أخرى –كتدابير دعم التغطية الصحفية المصمّمة لمواجهة المعلومات المُضلّلة– تكون بطبيعتها أكثر تحديداً في الزمن. ولابدّ من الإشارة هنا إلى أن التحديات والفرص المختلفة تندرج ضمن أطر زمنية مختلفة.

- أوجه التكامل: تكمّل أنواع الاستجابات للتضليل الأحد عشر المذكورة في هذه الدراسة بعضها البعض في جوانب كثيرة. ويمكن الاعتراف بها كمجموعة شاملة من التدخلات. فعلى سبيل المثال، كشف صحفيون في حالات كثيرة معلومات مضللة عبر الإنترنت لم تكن قد اكتشفتها (ولا اعترفت بها) شركات ومنصات التواصل عبر الإنترنت، الأمر الذي أتاح نشرها. وفي إطار الاستجابات، يجب أن تحظى الإجراءات التي تتخذها هذه الشركات بالاهتمام، لأن استخدام القوة والسياسات العامة والاهتمام بالجماهير هي من فئات الاستجابات التي لا تستطيع وحدها «حل» مشكلة التّضليل إذ لابدّ لها أن تتماشى مع إجراءات يتّخذها هذا القطاع لوقف نقل المعلومات المُضلّلة.

- التناقضات: هناك حالات يمكن أن يعمل فيها أحد أنواع الاستجابات ضد نوع آخر. ومن الأمثلة على ذلك اختلال التوازن الذي ينشأ متى تمّ التركيز بصورة مفرطة على اعتماد تنظيم من أعلى إلى أسفل، مع إهمال الحاجة إلى التمكين من القاعدة إلى القمة في الوقت نفسه. ومن الأمثلة الأخرى على ذلك مقاومة شركات ومنصات التواصل عبر الإنترنت لإزالة المُحتوى المرتبط بهجمات التّضليل التي تُشنّ على الصحفيين بحجّة «حُريّة الكلام». ويسلّط هذا الأمر الضوء على توتر يمكن أن يبرز متى استُخدمت «حُريّة الكلام» كحجة لتجنب المسؤولية عن الاستجابة السريعة والحازمة للتضليل الذي يقوّض حُريّة الصحافة (التي هي نتيجة طبيعية لحُريّة التّعبير) وسلامة الصحافة. وبعبارة أخرى، لا يشكّل الحفاظ على «حُريّة الكلام» من دون مراعاة حقوق حُريّة الصحافة (التي تشمل ضرورة حماية الصحفيين) والحفاظ عليها مقاربة مستدامة للتخفيف من التجاوزات في الاستجابات للتضليل.

- وقد يبرز توتّر آخر نتيجة وقوع الصحفيين في شباك نُصبت لمُروّجي التّضليل من خلال تجريم نشر معلومات مغلوطة أو توزيعها (على سبيل المثال عن طريق قوانين «الأخبار الزائفة»)، في وقت تكتسي فيه الصحافة أهمية كبرى لمواجهة التّضليل. ويمكن أيضاً الإشارة إلى ضرورة أن يتعايش المُحتوى المناهض للتضليل مع الصحافة المستقلة، لا أن يتنافس معها أو أن يكون على حسابها. لهذا السبب، لابد من مُواءمة التدخلات المختلفة، بدلاً من أن يسير كلّ منها في اتجاه منفصل.

- النوع الاجتماعي: تتغاضى استجابات كثيرة للتضليل عن النوع الاجتماعي، الأمر الذي قد يتجاهل الاختلافات الدقيقة في كيفية استهداف المُحتوى المغلوط للناس في كثير من الأحيان، بالإضافة إلى تجاهل الاختلافات في طريقة تجاوب الناس مع المُحتوى المعني. ومن المهمّ أيضاً الإشارة

إلى أنّ أنماط السلوك التي يعتمدها وكلاء التّضليل تشمل الهجمات الجنسانية على الإنترنت (التي تتراوح ما بين الإساءة والتهديد بالعنف الجنسي، مروراً بانتهاكات الأمن الرقمي والخصوصية). وتُسجّل أيضاً مسألة نفاذ النساء والفتيات إلى المعلومات، التي كثيراً ما تكون مقيدة في سياقات معينة والتي يهددها العنف المنزلي، مما قد يحد من إمكانية استفادتهنّ من جهود مُكافحة التّضليل.

- التركيبة السكّانية العمرية –لا سيّما الأطفال وكبار السن– لا تُراعَى في الكثير من الاستجابات لمشكلة التّضليل.

6.8 تقييم عام

تزدهر المعلومات المُضلّلة في غياب معلومات جديرة بالثقة يمكن التحقُّق منها. وبالمثل، يمكن أن تزدهر وسط كميات كبيرة من المُحتوى، عندما يجد الناس صعوبةً في التمييز بين المعلومات الموثوقة والمعلومات المُضلّلة، وبين الواقع التي تمّ التحقُّق منه وما هو غير ذلك. والمعلومات المُضلّلة تستغلّ حاجة الناس إلى إيجاد معنى للتطورات المعقدة، ولمخاوفهم وآمالهم وهُويّاتهم. هذا هو السبب الذي يدفع إلى اعتماد مقاربة متعددة الأوجه – وهي مقاربة تتجاوز أيضاً مجال التواصل والمُحتوى المتنازع عليه لتشمل خُطوات عملية مثل: الدمج والتضامن الاجتماعيَّين، وتعزيز الأخلاقيات والقيم على الصعيدَين الشخصي والمجتمعي، وترسيخ مبادئ بناء السّلام داخل المجتمعات عبر الإنترنت. وفي هذا الصدد، يتعيّن على أيّ استراتيجية متماسكة لمُكافحة تلوث المعلومات الاعتراف بقيمة اعتماد مقاربة شاملة وتحليلية إزاء المشكلة.

في هذا السياق الأوسع، من الواضح أن حُريّة التّعبير والنفاذ إلى المعلومات، والصحافة المستقلة – المدعومة كلها بوصول مفتوح وميسور التكلفة إلى الإنترنت– ليست فقط حقوقاً أساسية من حقوق الإنسان، بل هي أيضاً أجزاء أساسية من الترسانة التي نُناضلُ بها على الجبهة من أجل التصدّي للمعلومات المُضلّلة المزمنة –سواءً أكانت مرتبطة بوباء، أم بانتخابات، أم بقضايا استراتيجيّة مثل تغيُّر المناخ.

وتجدر الإشارة إلى أنّ مناهضة المعلومات المُضلّلة ليست دعوة لقمع تعددية المعلومات والرأي، ولا لإسكات النقاش النابض حول السياسات العامة. إنها معركة من أجل الوقائع، لأنه بدون معلومات قائمة على أدلة ومتاحة لكل شخص، لن يكون بالإمكان الوصول إلى معلومات موثوقة وذات مصداقية وقابلة للتحقق بشكل مستقل تدعم الديمقراطية وتجنّب تفاقم تداعيات الأزمات وصولاً إلى الأوبئة والجوائح المستعصية على الحل مثل جائحة فيروس كورونا.

7.8 الاستجابات للمعلومات المُضلّلة: إطار تقييم حُريّة التّعبير

صُممت أداة التقييم هذه المكوّنة من 23 خطوة لمساعدة الدول الأعضاء في اليونسكو على صياغة استجابات تشريعية وتنظيمية وسياستية للتصدي للتضليل؛ فيما تتيح، في الوقت نفسه، احترام حُريّة التّعبير والنفاذ إلى المعلومات وحقوق الخصوصية. ويمكن تطبيق هذه الأداة على التشريعات والسياسات المقترحة لإجراء تقييم –خطوة خطوة– لمدى مُلاءمتها للقوانين والقواعد الدولية لحقوق الإنسان.

1. هل شهدت الاستجابات انخراط أصحاب مصلحة متعددين (خاصةً منظمات المجتمع المدني والباحثين المتخصصين والخبراء في مجال حُريّة الصحافة) وتقديمهم إسهامات قبل الصياغة والتنفيذ؟ وفي حال الاستجابات التشريعية، هل توفّرت فرصة مناسبة للتداول قبل اعتمادها، وهل يمكن إجراء استعراض مستقلّ؟

2. هل تحدّد الاستجابات بوضوح وشفافية المشكلات التي يتعيّن مُعالجتها (مثل: التهوُّر أو النشاط الاحتيالي الفردي؛ وعمل شركات ومنصات التواصل عبر الإنترنت والمنظمات الإعلامية؛ والممارسات التي يقوم بها المسؤولون أو الجهات الفاعلة الأجنبية والتي تؤثر سلباً على الصحة العامة والسلامة أو النزاهة الانتخابية أو الحد من آثار تغيُّر المناخ وما إلى ذلك مثلاً)؟

3. هل تتضمّن الاستجابات تقييماً للأثر ينظر في العواقب المترتبة على الأُطُر الدولية لحقوق الإنسان التي تدعم حُريّة التّعبير أو حُريّة الصحافة أو الوصول إلى المعلومات أو الخصوصية؟

4. هل تؤثّر الاستجابات على حُريّة التّعبير والخصوصية والوصول إلى المعلومات أو تحدّ منها؟ إذا كان الأمر كذلك، وكانت الظروف التي تؤدي إلى الاستجابة مناسبة لمثل هذا التدخل (على سبيل المثال: جائحة فيروس كورونا)، فهل تم تحديد التداخل مع هذه الحقوق بشكل ضيق، وهل يُعتبر ضرورياً ومتناسباً ومحدداً زمنياً؟

5. هل تقيّد استجابة معينة الأعمال الصحفية، مثل الإبلاغ والنشر والتواصل بشأن سرية المصدر أو تضعها في خطر؟ وهل تحدّ من حقّ الوصول إلى المعلومات ذات المصلحة العامة؟ يمكن أن تشمل الاستجابات في هذه الفئة ما يلي: قوانين «الأخبار الزائفة»؛ والقيود المفروضة على حُريّة التنقّل والوصول إلى المعلومات بشكل عام، كما هي مطبّقة على موضوع معيّن (مثل: الإحصاءات الصحية والنفقات العامة)؛ واعتراض الاتصالات، والمراقبة المستهدفة أو الجماعية؛ والاحتفاظ بالبيانات وتسليمها. وإذا كانت هذه التدابير تمس بالفعل بهذه الوظائف الصحفية أو بمُساءلة المكلفين بواجبات أمام أصحاب الحقوق بوجه عام؛ يُرجى الرجوع إلى النقطة 4 أعلاه.

6. إذا كانت استجابة معينة تحدّ من أيّ من الحقوق المبينة في النقطة 4، هل توفّر استثناءات للأعمال الصحفية؟

7. هل يتم النظر في الاستجابات (كالاستجابات التثقيفية والمعيارية والقانونية وغيرها) معاً وبشكل كلي من حيث أدوارها المختلفة وأوجه تكاملها والتناقضات المحتملة؟

8. هل تُعدّ الاستجابات تقييدية في المقام الأول (مثل: القيود القانونية على التّضليل الانتخابي)، أم يتوفّر توازن مناسب مع التدابير التمكينية (مثل: زيادة تثقيف الناخبين والدراية الإعلامية والمعلوماتية)؟

9. في حين يمكن أن يكون تأثير المعلومات المُضلّلة والمغلوطة بنفس القدر من الخطورة، هل تعترف الاستجابات بالفرق في الدافع بين الجهات الفاعلة المتورطة في الأكاذيب المتعمّدة (المعلومات المُضلّلة) والجهات الضالعة في الأكاذيب غير المتعمّدة (المعلومات المغلوطة)، وهل تُصمم الإجراءات وفقاً لذلك؟

10. هل تمزج الاستجابات بين مُحتوى المعلومات المُضلّلة ومُحتوى خطاب الكراهية أو تساويهما ببعضهما البعض (على الرغم من أنّ المعايير الدولية تبرّر التدخلات القوية للحد من المُحتوى الأخير، في حين لا تُستبعد الأكاذيب في حد ذاتها من حُريّة التّعبير)؟

11. هل يستطيع الصحفيُّون والجهات الفاعلة السياسية والمدافعون عن حقوق الإنسان الحصول على حماية قضائية فعّالة من المعلومات المُضلّلة و/أو المُحتوى الذي يحثّ على الكراهية ويُحرّض على العداء والعنف والتمييز، ويهدف إلى تخويفهم؟

12. هل تترافق الاستجابات القانونية مع المشورة والتدريب ليتم تنفيذها من قبل السُّلطات المعنية بإنفاذ القانون والمُدَّعين العامين والقُضاة بغية حماية الحق الأساسي في حُريّة التّعبير وتبعات تقييد هذا الحق؟

13. هل بالإمكان تقييم الاستجابة بشفافية، وهل هناك عملية لرصد التبعات على حُريّة التّعبير وتقييمها بشكل منهجي؟

14. هل تخضع الاستجابات لتدابير رقابة ومُساءلة، بما في ذلك نظم المُراجعة والمُساءلة (كالتقارير المرفوعة إلى الجمهور والبرلمانيين وأصحاب مصلحة معينين)؟

15. هل يمكن الطعن في استجابة معينة أو هل يمكن التراجع عنها إذا تبيّن أنّ آثارها السلبية على حُريّة التّعبير والنفاذ إلى المعلومات وحقوق الخصوصية (التي هي في حد ذاتها ترياق للتضليل) تتخطّى منافعها؟

16. هل وُضعت التدابير المتعلقة بشركات ومنصات التواصل عبر الإنترنت مع إيلاء الاعتبار الواجب لمشاركة أصحاب مصلحة متعددين ولصالح تعزيز الشفافية والمُساءلة، مع تجنّب خصخصة الرقابة؟

17. هل يتوفّر تقييم (مُستنير، بمشورة الخبراء) للإمكانات التي تنطوي عليها الاستجابات التكنولوجية التي تتناول المعلومات المُضلّلة وحدودها في آنٍ واحد (مع الحفاظ على حُريّة التّعبير والخصوصية)؟ هل تتوفّر توقعات غير واقعية بشأن دور التكنولوجيا؟

18. هل تشارك الجهات الفاعلة في المجتمع المدني (بما في ذلك المنظمات غير الحكومية والباحثون ووسائل الإعلام) كشركاء مستقلين في مُكافحة المعلومات المُضلّلة؟

19. هل تدعم الاستجابات إنتاج المعلومات وتأمينها ونشرها –بما في ذلك المعلومات المحلية ومتعددة اللُّغات– كبديل موثوق عن التّضليل؟ من الأمثلة على ذلك تقديم إعانات للصحافة

الاستقصائية التي تنظر في المعلومات المُضلّلة، ودعم الإذاعة المجتمعية ووسائل الإعلام الناطقة بلغة الأقليات.

20. هل تشمل الاستجابات تقديم الدعم للمؤسسات (مثل: رسائل الخدمة العامة وإعلاناتها؛ والمدارس) لإتاحة المجال للتصدي للتضليل؟ يمكن أن يشمل ذلك تدخلات من قبيل الاستثمار في المشروعات والبرامج المصمّمة خصيصاً للمساعدة على «تحصين» المجتمعات بشكل عام ضد التّضليل من خلال برامج الدراية الإعلامية والمعلوماتية.

21. هل تعظّم الاستجابات انفتاح البيانات التي تحتفظ بها سلطات الدولة وتوفّرها، مع إيلاء الاعتبار اللازم لحماية الخصوصية الشخصية، في إطار حق النفاذ إلى المعلومات والإجراءات الرسمية التي تهدف إلى استباق الشائعات والاضطلاع بالأبحاث والتقارير المستندة إلى الوقائع؟

22. هل تراعي الاستجابات النوع الاجتماعي وأوجه الضعف المحددة (مثل: الشباب والمُسنين) الناجمة عن التعرض للمعلومات المُضلّلة وتوزيعها وآثارها؟

23. إن تم اعتماد تدابير الاستجابة رداً على مشكلة مُلحّة، أو جرى تصميمها لتأتي بتأثير قصير الأجل (مثل التدخلات الحساسة من ناحية الوقت التي ترتبط بالانتخابات)، فهل تصحبها مبادرات أو برامج أو حملات تهدف إلى إحداث التغيير وترسيخه على المدى المتوسط والبعيد؟

8.8 توصيات شاملة للعمل

تستند التوصيات الواردة أدناه إلى التوصيات الخاصة بكل فصل، والتي تتمحور حول أنواع معينة من المعلومات المُضلّلة. وتجمع هذه التوصيات النقاط الرئيسية من الفصول من أجل وضع قائمة كاملة بالخيارات المتوفّرة أمام كل مجموعة من مجموعات أصحاب المصلحة في ما يتعلق بأنواع المعلومات المُضلّلة. ويتيح هذا الأمر توفير نظرة شاملة سهلة الاستخدام عن المجموعة الكلية من الإجراءات التي تستطيع كل مجموعة من أصحاب المصلحة النظر في إمكانية اتخاذها من أجل تعزيز فعالية استجاباتها وتقوية الأبعاد المتعلقة بحُريّة التّعبير فيها. وفي الوقت نفسه، يتم الإقرار بأنّ الشراكات داخل كل مجموعة من أصحاب المصلحة وعبرها ضرورية لتحقيق النجاح.

توصيات مُتقاطعة تستهدف الجهات الفاعلة كلّها:

- التشجيع على تعزيز المجموعة المتنوعة من الاستجابات للمعلومات المُضلّلة، والحرص على امتثالها كلّها للمعايير الدولية لحقوق الإنسان.
- تيسير التعاون العالمي المُعد له بين أصحاب المصلحة المتعددين وتبادل الممارسات الجيدة عبر القارات والدول وتشجيعهما؛ من أجل التنفيذ الفعّال للتدابير الشاملة الرامية إلى التصدي للتضليل عبر الإنترنت.
- تشجيع المانحين على الاستثمار على وجه التحديد في التدابير المضادة للتضليل التي تعزّز الدراية

الإعلامية والمعلوماتية، وحُريّة التّعبير، والصحافة المستقلة، والتطوير الإعلامي.

- زيادة الشفافية الرسمية والكشف الاستباقي عن المعلومات والبيانات الرسمية، ورصد هذا الأداء بما يتماشى مع الحق في الوصول إلى المعلومات والمؤشر 16.10.2 الخاص بأهداف التنمية المستدامة الذي يُقيّم اعتماد الضمانات الدستورية والقانونية و/أو السياساتية لوصول الجمهور إلى المعلومات وتنفيذها.
- تعزيز الحفاظ على الخصوصية، والوصول العادل إلى البيانات الرئيسية من شركات ومنصات التواصل عبر الإنترنت لإتاحة المجال أمام إجراء تحليل مستقل لحدوث التّضليل عبر الإنترنت ونشر المعلومات المُضلّلة وتأثيرها على المواطنين حول العالم، لا سيّما في سياق الانتخابات والصحة العامة والكوارث الطبيعية.
- الاستثمار في أبحاث مستقلة تركّز على الطبيعة سريعة التغيّر للاستجابات للتضليل وعلى نطاقها، وعلى ضرورة مواجهة تحديات دراسة المنصات الاجتماعية الجديدة؛ سريعة التطور، بما في ذلك المنصات التي يُنظر إليها بشكل رئيسي على أنها فسح ترفيهية واجتماعية (مثل: تيك توك).

توصيات للجنة النطاق الواسع:

- مواصلة رصد آثار الاستجابات للتضليل على أُطُر حقوق الإنسان وقياسها وتقييمها، بما في ذلك استخدام إطار التقييم الوارد أعلاه.
- تشجيع أعضاء اللجنة الذين يمثّلون شركات ومنصات التواصل عبر الإنترنت على التأكّد من أنّ الاستجابات التي يُطلقونها شفافة وقابلة للقياس بشكل مناسب، وأن تنفيذها يحصل على نطاق عالمي حقيقي.
- تشجيع الشركات الأعضاء على النظر في الاستجابات السريعة والحاسمة للتضليل السياسي والانتخابي، كما حدث في مجال المعلومات المُضلّلة ذات الصلة بجائحة فيروس كورونا، مع إيلاء الاعتبار الواجب للفرق بين هذين المجالين.
- تشجيع الأعضاء على إدماج هذه الدراسة في أنشطتهم ولفت انتباه أصحاب المصلحة إليها.

توصيات للمنظمات الحكومية الدولية والمنظمات الدولية الأُخرى، حسب الاقتضاء:

- زيادة المساعدة التقنية المقدّمة إلى الدول الأعضاء بناءً على طلبها من أجل المساعدة في وضع الأُطر والسياسات التنظيمية، تماشياً مع المعايير الدولية لحُريّة التّعبير واحترام الخصوصية، لمعالجة التّضليل على الإنترنت. ويمكن أن ينطوي ذلك على التشجيع على اعتماد إطار تقييم الاستجابات للتضليل الذي يتألف من 23 خطوة والذي وُضع لأغراض هذه الدراسة.
- بالنسبة إلى اليونسكو على وجه التحديد التي تتمتّع بولاية تتمحور حول حُريّة التّعبير؛ تكثيف العمل الجاري بشأن المعلومات المُضلّلة بالشراكة مع منظمات الأمم المتحدة الأخرى ومجموعة من الجهات الفاعلة المنخرطة في هذا المجال.

- الاستثمار في البحث عن آثار الاستجابات للتضليل على أُطُر حقوق الإنسان ورصدها وقياسها وتقييمها، بما في ذلك عبر استخدام إطار التقييم المعروض هنا.
- العمل مع الدول والمنظمات غير الحكومية على إطلاق مبادرات في مجال الدراية الإعلامية والمعلوماتية تستهدف الفئات المُستضعفة المحتملة.
- النظر في إمكانية عقد مؤتمرات متعددة اللغات، وتبادل المعرفة، وورش عمل تركز على الدراية الإعلامية والمعلوماتية كاستجابة للتضليل.
- زيادة العمل في مجال الدراية الإعلامية والمعلوماتية وتدريب الصحفيين باعتبار أنّ ذلك يُشكّل استجابة هامة للتضليل.
- زيادة الدعم المقدّم إلى المؤسسات الإعلامية في البلدان النامية، بما في ذلك من خلال برنامج اليونسكو الدولي لتطوير الاتصالات، لتمكينها من مواصلة إنتاج الصحافة ذات المصلحة العامة وتجهيزها لمُكافحة التّضليل.
- دعم الاستجابات التي تُراعي الاعتبارات الجنسانية في مجال المعلومات المُضلّلة.
- تشجيع المانحين على الاستثمار على وجه التحديد في التدابير المضادة للتضليل؛ التي تعزّز عملية التحقُّق من الوقائع المستقلة، والدراية الإعلامية والمعلوماتية، وحُريّة التّعبير، والصحافة المستقلة، والتطوير الإعلامي.

توصيات للدول الفردية:

- رفض مُمارسة تداول المعلومات المُضلّلة بأيّ شكل، بما في ذلك الالتزام بعدم الانخراط في التلاعب بالرأي العام بشكل مباشر أو غير مباشر- مثلاً، من خلال «عمليات التأثير» التي ينتجها مشغلون من الأطراف الثالثة، كشركات العلاقات العامة التي تستعمل «الدعاية المظلمة».
- استعراض استجاباتها للتضليل وتكييفها باستخدام الإطار المؤلَّف من 23 خطوة لتقييم القوانين والسياسات التي وُضعت كمخرج لهذه الدراسة بهدف الامتثال للمعايير الدولية لحقوق الإنسان (لا سيما حُريّة التّعبير، بما في ذلك النفاذ إلى المعلومات، وحقوق الخصوصية)، مع وضع ترتيبات لرصد استجاباتها وتقييمها في الوقت نفسه.
- زيادة الشفافية والإفصاح الاستباقي عن المعلومات والبيانات الرسمية، ورصد هذا الأداء بما يتماشى مع الحق في الحصول على المعلومات، والمؤشر 16.10.2 الخاص بأهداف التنمية المستدامة الذي يقيّم اعتماد الضمانات الدستورية والقانونية و/أو السياساتية لنفاذ الجمهور إلى المعلومات، وتنفيذها.
- تعزيز الربط ميسور التكلفة للجميع بتكنولوجيا المعلومات والاتصالات بما يتماشى مع مفهوم اليونسكو «لعالمية الإنترنت ومبادئ ROAM الأربعة (أي الحقوق والانفتاح وإمكانية الوصول ومشاركة أصحاب المصلحة المتعددين).

- دعم إنشاء مجالس و/أو مرافق وطنية شفّافة للتظلُّمات؛ تخصّ وسائل التواصل الاجتماعي من أجل تمكين المستخدمين من اللُّجوء إلى التحكيم المستقل والطعون ضد خُطوات مراقبة المُحتوى التي تُعتبر غير عادلة.
- دعم الاستثمار في تعزيز وسائل الإعلام المستقلة، بما في ذلك وسائل الإعلام المجتمعية والعامة، نظراً إلى أنّ الآثار الاقتصادية لجائحة فيروس كورونا؛ تُهدّد الاستدامة الصحفية في جميع أنحاء العالم.
- تخصيص التمويل والدعم للدراية الإعلامية والمعلوماتية التي تركّز على مُكافحة المعلومات المُضلّلة، لا سيما من خلال التدخلات التثقيفية التي تستهدف الأطفال والشباب والمُسنّين والفئات المُستضعفة المحتملة.
- العمل مع شركات ومنصات التواصل عبر الإنترنت لتبادل البيانات بصورة آمنة تحفظ الخصوصية، ولتسهيل وصول الصحفيين والباحثين الأكاديميين والباحثين في المنظمات غير الحكومية إلى بيانات وسائل التواصل الاجتماعي عند الاقتضاء، لإتاحة المجال أمام إجراء تحقيقات شاملة والحفاظ على البيانات الهامة تاريخياً (لا سيما تلك المتصلة بالانتخابات والأوبئة وغيرها من الموضوعات الهامة).
- تجنّب تجريم التّضليل للحرص على عدم وقوع الصحافة المشروعة وغيرها من المعلومات ذات المصلحة العامة في شباك قوانين «الأخبار الزائفة».. إلخ.
- التأكُّد من أنّ أيّ تشريع أو لائحة تنظيمية تستجيب لمشكلات المعلومات المُضلّلة، مثل أزمة التّضليل في ظل جائحة فيروس كورونا، ضرورية ومتناسبة ومحددة زمنياً.
- تطوير آليات للرقابة والتقييم المستقلّين للتبعات على حُريّة التّعبير ولفعالية التشريعات ذات الصلة «بالأخبار الزائفة»، فضلاً عن السياسات الوطنية والمبادرات المعيارية الأخرى ذات الصلة.
- الحرص على مُراعاة الاعتبارات الجنسانية في استراتيجياتها والاستجابات العامة للتضليل.
- التشجيع على الأخذ بالتوصيات الواردة أدناه، التي تخصّ الأحزاب السياسية والجهات الفاعلة في مجال الانتخابات والحملات الانتخابية.

توصيات للهيئات الناظمة للانتخابات والسُّلطات الوطنية:

- تعزيز التدابير القانونية المتعلقة بحماية الخصوصية، وحُريّة التّعبير، والإعلان السياسي من أجل توفير حماية أفضل من المعلومات المُضلّلة الانتخابية.
- تحسين شفافية الإعلانات الانتخابية التي تُطلقها الأحزاب السياسية والمرشحون والمنظمات التابعة لهم من خلال اشتراط توفّر قواعد بيانات إعلانية شاملة ومتاحة علناً والكشف عن الإنفاق من قبل الأحزاب السياسية ومجموعات الدعم.
- إرساء تعاون فعّال مع شركات التواصل عبر الإنترنت لمُراقبة التهديدات التي تحدق بنزاهة الانتخابات والتصدي لها.

- السعي إلى وضع استجابات تشمل أصحاب مصلحة متعددين، لا سيما المجتمع المدني، والترويج لها.
- تثقيف المواطنين وتمكينهم من الكشف عن المعلومات المُضلّلة والإبلاغ عنها خلال الانتخابات.
- تطوير إلمام الناخبين من خلال الربط بين التربية المدنية والتثقيف في مجال المواطنة الرقمية والدراية الإعلامية والمعلوماتية.
- العمل مع الصحفيين والباحثين على التحقُّق من الوقائع وإجراء تحقيقات حول شبكات المعلومات المُضلّلة الانتخابية ومنتجي «الدعاية المظلمة».

توصيات للأحزاب السياسية وغيرها من الجهات الفاعلة السياسية:

- التحدّث علناً عن المخاطر التي تطرحها الجهات الفاعلة السياسية كمصادر للمعلومات المُضلّلة وعناصر مضخّمة لها، والعمل على تحسين نوعية النظام البيئي للمعلومات وزيادة الثقة في المؤسسات الديمقراطية.
- الامتناع عن استخدام تكتيكات التّضليل في الحملات السياسية، بما في ذلك الأدوات السرية للتلاعب بالرأي العام وشركات العلاقات العامة المعنية بـ«الدعاية المظلمة».
- النظر في إمكانية متابعة الأحزاب السياسية خلال الانتخابات الأخيرة التي تعهّد فيها المتنافسون بتجنب التّضليل[487].
- الالتزام بالشفافية والمُساءلة في مجال التدقيق الذي تُجريه الجهات الفاعلة الصحفية الناقدة وغيرها من الآليات الداعمة للمجتمعات المفتوحة، وإدانة التهديدات الموجهة ضد الصحفيين، بما في ذلك استخدام المعلومات المُضلّلة كسلاح ضد وسائل الإعلام الإخبارية.
- إخضاع إعلاناتها السياسية عبر الإنترنت لعمليات التحقُّق من الوقائع المستقلة.

توصيات لسُلطات إنفاذ القانون والقضاء:

- الحرص على إدراك موظفي إنفاذ القانون لأهمية حُريّة التّعبير وحقوق الخصوصية، بما في ذلك الحماية الواجب منحها للصحفيين الذين ينشرون معلومات يمكن التحقُّق منها من أجل المصلحة العامة، وتفادي الإجراءات التعسفية المرتبطة بأيّ قوانين تجرّم التّضليل.
- بالنسبة إلى القضاة وغيرهم من الجهات القضائية الفاعلة: التنبُّه الشديد عند استعراض القوانين والقضايا المتعلقة بالتدابير الآيلة إلى مُكافحة المعلومات المُضلّلة، مثل التجريم، من أجل المساعدة على ضمان الاحترام الكامل للمعايير الدولية المتعلقة بحُريّة التّعبير والخصوصية في إطار تلك التدابير.

---

[487] اتفقت الأحزاب السياسية في أوروغواي في عام 2019 على ميثاق للامتناع عن التّضليل، وتفادي الإجراءات أو التعبيرات التي تستخدم نبرات متشددة ضد الخصوم، وإنشاء آلية تشاورية متى برزت تهديدات أو تحديات عند تنفيذ اتفاقها:
https://www.undp.org/content/dam/uruguay/docs/GD/undp-uy-pacto-etico-definformacion.pdf/
وفي ألمانيا، تعهّدت الأحزاب السياسية بتفادي «روبوتات الويب» و«الاستهداف الدقيق» عبر وسائل التواصل الاجتماعي:
https://www.bpb.de/gesellschaft/digitales/digitale-desinformation/290568/relevanz-und-regulierung-von-social-bots

توصيات لشركات ومنصات التواصل عبر الإنترنت:

- تكثيف مشاركة أصحاب المصلحة المتعددين وتعزيز الشفافية في سياساتها بوجه عام وتطبيقها، بما في ذلك استجابتها للتضليل.
- تنفيذ استجاباتها على نطاق عالمي، بدلاً من أن تقتصر على بلدان معينة، وضمان التغطية بجميع اللُّغات المعنيّة.
- توفير المزيد من الدعم المالي لما يلي: الشبكات المستقلة للتحقق من الوقائع، والصحافة المستقلة (لا سيّما تلك التي تُركّز على التحقيقات التي تستهدف مُحتوى المعلومات المُضلّلة وشبكاتها، فضلاً عن الوكالات الإخبارية المحلية الهشة بشكل خاص)، ومبادرات الدراية الإعلامية والمعلوماتية المُقدَّمة/المنفّذة بشكل مستقل.
- تفادي التدخلات التي تبدو مصممة في المقام الأول كتدريبات في العلاقات العامة أو في إدارة العلامات التجارية، وتقديم إسهامات من دون فرض أي قيود، وتحسين الشفافية المتعلقة بالتمويل.
- دعم الصناديق ذات الإدارة المستقلة المخصّصة للبحث في حالات التّضليل، وتأثيرها، والاستجابات لها، بما في ذلك التقييمات المستقلة لفعالية الاستجابات التي تضعها الشركات للتضليل. الحرص على تنوع المستفيدين من التمويل وعلى شفافية أساليب البحث والنتائج.
- العمل معاً، من خلال إطار خاص بحقوق الإنسان، للتعامل مع المعلومات المُضلّلة العابرة للمنصات، من أجل تحسين القدرات التكنولوجية للكشف عن مشكلات المُحتوى الكاذب والمُضلّل والحد منه بفعالية أكبر، وتبادل البيانات حول هذا الموضوع.
- تطوير استجابات تستند إلى تنظيم المُحتوى لضمان وصول المستخدمين بسهولة إلى الصحافة بصفتها معلومات يمكن التحقُّق منها ويتم مشاركتها في سبيل المصلحة العامة، مع إعطاء الأولوية للمنظمات الإخبارية التي تمارس الصحافة المستقلة الأخلاقية والناقدة.
- العمل على إبراز المُحتوى الإخباري الموثوق به، ودفع تعويضات مالية لصنّاع الأخبار الذين يفيد محتواهم أعمالها، خاصّةً وأن الكثير من المؤسسات الإخبارية أزالت نظام الدفع والحواجز الأخرى التي تعترض الوصول إلى المُحتوى خلال جائحة فيروس كورونا كتدبير مضاد للتضليل.
- تجنّب الاعتماد المفرط على الميكنة لمراقبة المُحتوى، والاعتراف بالحاجة إلى توسيع قدرة المراجعة البشرية وسبل الإنصاف، ورصد هذه المسائل بشفافية.
- توفير الأجر والتدريب والدعم النفسي بصورة مناسبة للأشخاص الذين يعملون على مراقبة المُحتوى.
- الاعتراف بأنه إذا أمكنت معالجة المعلومات الصحية المُضلّلة والمغلوطة بسرعة خلال جائحة ما نظراً إلى الخطر الجسيم الذي تُشكّله على الصحة العامة، فلابد أيضاً من اتخاذ إجراءات ضد التّضليل السياسي - خاصةً عند تقاطعه مع خطاب الكراهية - عندما يصبح هو أيضاً مُهدّداً للحياة. وينطبق الأمر نفسه على المعلومات المُضلّلة المتصلة بقضية تغيُّر المناخ.
- الإقرار بأن حُريّة الصحافة وسلامة الصحفيين مكوّنان حاسمان من الحق في حُريّة التّعبير، مما يعني أنه لا يمكن التسامح مع العنف عبر الإنترنت الذي يستهدف الصحفيين (وهو سمة متكررة لحملات التّضليل).

- اعتماد أسلوب التحقُّق من الوقائع عند التعامل مع كل المحتويات السياسية (بما في ذلك الإعلانات والرأي القائم على الوقائع والكلام المباشر) التي ينشرها السياسيُّون والأحزاب السياسية والمنتسبون إليها والجهات الفاعلة السياسية الأخرى.

- إصدار تقارير مفصلة ومتكررة عن الشفافية العامة، بما في ذلك معلومات محددة عن أصول المعلومات المُضلّلة ونطاقها وآرائها وتدفقها وأنواعها، وإزالة المعلومات المُضلّلة، وحجب الربح المالي عن مُحتوى التّضليل، وتعليق الحسابات التي تنشر معلومات مُضلّلة، فضلاً عن توفير معلومات عن الخُطوات الأخرى لتنظيم المُحتوى، كالتصنيف والتظلُّمات.

توصيات لقطاع الإعلام:

- مضاعفة الجهود كمستجيب مهني في خط المواجهة مع التّضليل، من خلال زيادة الاستثمار في التحقُّق من الوقائع، وكشف التّضليل، والتحقيقات في المعلومات المُضلّلة، وضمان وجود خط استجواب قوي حول الاستجابات للمعلومات المُضلّلة، وتعزيز المُساءلة والشفافية في ما يتعلق بالجهات الفاعلة السياسية والدول والمؤسسات وقطاع الشركات.

- الإبلاغ عن تبعات الاستجابات للتضليل على حقوق الإنسان، بما في ذلك الاستجابات التي تؤثر على حُريّة التّعبير والنفاذ إلى المعلومات، فضلاً عن حقوق الخصوصية.

- النظر في دحض الخرافات وإقامة تعاون استقصائي يطال المعلومات المُضلّلة مع الوكالات الإخبارية الأخرى والجماهير، بما في ذلك على الصعيد الدولي. ويمكن أيضاً أن تحقّق الشراكات مع الباحثين ومنظمات المجتمع المدني النجاح.

- تركيز جهود مبتكرة على مُكافحة المعلومات المُضلّلة من خلال أشكال القصص الجذابة التي يَسْهُل الوصول إليها، كالمخططات البيانية والبودكاست (المُدونة الصوتية) إلى جانب التحقيقات التعاونية القائمة على البيانات.

- مراعاة الخبرات المتوفّرة في مجموعة من البلدان النامية لفائدة تغطية المعلومات المُضلّلة والاستجابات لها.

- ضمان استعداد الموظفين لمواجهة مخاطر السلامة المرتبطة بالإبلاغ عن المعلومات المُضلّلة، مثل زيادة التهديدات الأمنية، والإساءة على الإنترنت، والاعتداءات البدنية، وضمان مراعاة الاعتبارات الجنسانية في الاستجابة لهذه المخاطر.

- تغطية المسائل المتعلقة بشفافية المؤسسات والأفراد المشاركين في التحقُّق من الوقائع و/أو تقييم مصداقية مصادر المعلومات، وبمُساءلتهم واستقلاليتهم.

توصيات للمجتمع المدني:

- دعوة أقوى لاعتماد استجابات للتضليل تتفق مع المعايير الدولية لحقوق الإنسان.

- إقامة شراكة مع الصحفيين والوكالات الإخبارية والباحثين في إطار مشروعات استقصاء ورصد تتناول التّضليل والاستجابات له.

- تعزيز نشر مشروعات الدراية الإعلامية والمعلوماتية والبرامج التي تدعم الصحافة المستقلة.
- النظر في إمكانية وضع برامج تستهدف الأطفال والمُسنّين الذين لا يحصلون على خدمات كافية من خلال حملات الدراية الإعلامية والمعلوماتية، والذين هم بالتالي أكثر عُرضةً للاستغلال من قبل وكلاء التّضليل.
- إنتاج مُحتوى مُضاد وإطلاق حملة ضد المعلومات المُضلّلة.

توصيات للباحثين:

- تعزيز أجندة التحقيقات العلمية الخاصة بهم للتركيز على المعلومات المُضلّلة والاستجابات لها، وآثار هذه الاستجابات.
- دراسة الأشكال التي لم تُغطّها الأبحاث بشكل كافٍ، مثل: الألعاب التفاعلية، حيث يستطيع التّضليل والتدابير المضادة استهداف الشباب بشكل فعالٍ.
- الشروع في عمل مشروعات بحثية تُركّز على العمل التشاركي وتستجيب للحوادث الحرجة المرتبطة بالتّضليل، وتستطيع أيضاً أن توفر المعرفة العاجلة.
- التعاون مع الصحفيين والوكالات الإخبارية ومجموعات المجتمع المدني في المشروعات التي تُساعد على الكشف عن المعلومات المُضلّلة ومكافحتها، إلى جانب تدريبات المُتابعة والرصد والتقييم التي تركّز على الاستجابات لها.
- دراسة حملات التّضليل العابرة للمنصات لاكتساب منظور أكثر شموليةً حول المشكلة والاستجابات لها.
- متابعة الرصد المستقل والطُّولي والكمي والنوعي للاستجابات للتضليل التي تنفذها شركات ومنصات التواصل عبر الإنترنت، وتقييمها.
- تطوير أدوات تكنولوجية جديدة لمساعدة الصحفيين وغيرهم من المهنيين في مجال التحقُّق على كشف المعلومات المُضلّلة وتحليلها، والحرص على إدراج حُريّة التّعبير والخصوصية والشفافية الخوارزمية وإمكانية الوصول في صميم هذه الأدوات.
- استخدام التصنيف وإطار التقييم اللَّذين تم تطويرهما من خلال هذه الدراسة لتحليل مختلف أنواع الاستجابات للتضليل عند ظهورها وتطوُّرها، وتقييم فعاليتها وآثارها مع الإشارة بشكل خاص إلى التحديات المتعلقة بحُريّة التّعبير.
- الاستثمار في دراسات تُعالج آثار التّضليل على المستخدمين، بما في ذلك التحقيقات في علم السُّلوك التي تبني المعرفة حول ما يُحفّز الناس على مشاركة المعلومات المُضلّلة و/أو منحها المصداقية من دون قصد.
- إعطاء الأولوية للدراسات التي تستهدف سُلوك المستخدمين لفائدة استعمال المعلومات الموثوقة والمُتحقَّق منها وإعادة توزيعها، كتلك التي ينتجها ناشرو الأخبار والصحفيون المستقلون.
- إبراز دور الخبيرات كأسلوب لمعالجة أوجه انعدام المساواة بين الجنسَين في النقاشات الدولية المتعلقة بالتّضليل.





الفصل التاسع:
قائمة المصادر
التي تمت مراجعتها

Aaron, C. (2020). 'Journalism Needs a Stimulus. Here's What it Should Look Like'. *CJR*, 24 March 2020. https://www.cjr.org/analysis/journalism-stimulus.php

Abellán, L. (2019). 'El Gobierno activa una unidad contra la desinformación ante las elecciones'. *El Pais*, 11 March 2019. https://elpais.com/politica/2019/03/10/actualidad/1552243571_703630.html

Adair, B. (2018). 'Beyond the Truth-O-Meter'. *Columbia Journalism Review*, 24 July 2018. https://www.cjr.org/first_person/beyond-the-truth-o-meter.php

Adobe. (2019). 'Introducing the Content Authenticity Initiative'. 04 November 2019. https://theblog.adobe.com/content-authenticity-initiative/

Adhikari, R. (2020). ''Black PR' Firms Line Their Pockets by Spreading Misinformation'. *Ecommerce Times*, 09 January 2020. https://www.ecommercetimes.com/story/86444.html

AEC. (Australian Electoral Commission). (2019a). 'Electoral Integrity: 2019 Federal Election'. https://www.aec.gov.au/elections/electoral-advertising/electoral-integrity.htm

AEC. (Australian Electoral Commission). (2019b). 'Stop and Consider'. https://www.aec.gov.au/Elections/electoral-advertising/stopandconsider.htm

AFP. (2019a). 'Sri Lanka proposes new law on fake news after Easter attacks'. *France 24*, 05 June 2019. https://www.france24.com/en/20190605-sri-lanka-proposes-new-law-fake-news-after-easter-attacks

AFP. (2019b). 'Nine arrested in Thailand for posting election 'fake news'. *France 24*, 28 March 2019. https://www.france24.com/en/20190328-nine-arrested-thailand-posting-election-fake-news

AFP. (2019c). 'Côte d'Ivoire: le député pro-Soro Lobognon condamné à un an de prison pour un tweet'. L*a Libre Afrique*, 30 January 2019. https://afrique.lalibre.be/31666/cote-divoire-le-depute-pro-soro-lobognon-condamne-a-un-an-de-prison-pour-un-tweet/

AFP. (2020). 'Legal challenge to Singapore misinformation law rejected'. 05 February 2020. http://theindependent.sg/legal-challenge-to-singapore-misinformation-law-rejected/

Africa Check, Chequeado, Full Fact. (2020). 'What is the impact of fact checkers' work. 13 March 2020. https://fullfact.org/blog/2020/mar/long-game-impact-fact-checkers/

Agrawal, A. (2019). 'Intermediary Guidelines to be notified by Jan 15, 2020, MeitY tells Supreme Court', *Medianama*, 21 October 2019. https://www.medianama.com/2019/10/223-intermediary-guidelines-to-be-notified-by-jan-15-2020-meity-tells-supreme-court/

Agarwal, S, Farid, H., Gu, Y., He, M., Nagano, K. & Li, H. (2019). 'Protecting World Leaders Against Deep Fakes'. *CVPR Workshop Paper*. http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.pdf

Ainge Roy, E. (2019). 'New Zealand bans foreign political donations amid interference concerns'. *The Guardian*, 03 December 2019. https://www.theguardian.com/world/2019/dec/03/new-zealand-bans-foreign-political-donations-amid-interference-concerns

Alam, J. (2018). 'Journalists Slam Pending Bangladesh Digital Security Law', *Associated Press*, 28 September 2018. https://apnews.com/6c7914d0733249aabacd50a3c40542fe/Journalists-slam-pending-Bangladesh-digital-security-law

Alaphilippe, A., De Marez, L., Gerlache, A., Lievens, E., Pauwels, T., Picone, I. & Rouvroy, A. (2018a) 'Verslag van de Belgische Expertengroep inzake Fake News en Desinformatie'. 18 July 2018. https://www.dropbox.com/s/99iza9kmbwjbels/20180718_rapport_onlinedesinformatieNL.pdf?dl=0

Alaphilippe, A., Bontcheva, K., Gizikis, A. Hanot, C., (2018b). 'Automated tackling of disinformation'. *European Parliamentary Research Service. Scientific Foresight Unit* (*STOA*). PE 624.278, March 2019. https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624278/EPRS_STU(2019)624278_EN.pdf

AlBadawy, E. A. Lyu, S. & Farid, H. (2019). 'Detecting AI-Synthesized Speech Using Bispectral Analysis'. *ResearchGate*, June 2019. https://www.researchgate.net/publication/333393640_Detecting_AI-Synthesized_Speech_Using_Bispectral_Analysis

Al Busaidi, A. S. H. (2019). 'Tackling fake news head-on in Oman'. *Times of Oman*, 28 July 2019. https://timesofoman.com/article/1681323/Oman/Tackling-fake-news-head-on-in-Oman

Alexander, J. (2020). 'YouTube is demonetizing videos about coronavirus, and creators are mad'. *The Verge*, 04 March 2020. https://www.theverge.com/2020/3/4/21164553/youtube-coronavirus-demonetization-sensitive-subjects-advertising-guidelines-revenue

Al Jazeera. (2018). 'Bangladesh Shuts Down Mobile Internet in Lead Up to Election Day'. 28 December 2018. https://www.aljazeera.com/news/2018/12/bangladesh-shuts-mobile-internet-lead-election-day-181229111353218.html

Allan, R. (2018). 'Hard Questions: Where Do We Draw the Line on Free Expression?' *Facebook*, 09 August 2018. https://about.fb.com/news/2018/08/hard-questions-free-expression/

Allcott, H, Gentzkow, M. (2017). 'Social Media and Fake News in the 2016 Election'. *Journal of Economic Perspectives* - Volume 31, Number 2, Spring 2017.

Allcott, H. Gentzkow, M. & Yu, C. (2018). 'Trends in the Diffusion of Misinformation on Social Media'. *Stanford, Institute for Economic Policy Research* (*SIEPR*). https://siepr.stanford.edu/system/files/publications/18-029.pdf

Allen, V. (2020). 'What Does Facebook's New Oversight Board Mean for Conservative Posts?'. *The Daily Signal*, 14 May 2020. https://www.dailysignal.com/2020/05/14/what-does-facebooks-new-oversight-board-mean-for-conservative-posts/

Allyn, B. (2020). 'Twitter Flags President Trump's Tweet For The 1st Time'. *NPR*, 26 May 2020. https://www.npr.org/2020/05/26/862838470/twitter-flags-president-trumps-tweet-for-the-first-time

Alves, L. (2018). 'Brazil Preparing to Fight Fake News During October's Elections', *The Rio Times*, 29 June 2018. https://riotimesonline.com/brazil-news/rio-politics/brazil-preparing-to-fight-fake-news-during-octobers-elections/

ANCIR. (n.d). 'Manufacturing Divides: The Gupta-linked Radical Economic Transformation (RET) media network'. https://s3-eu-west-1.amazonaws.com/s3.sourceafrica.net/documents/118115/Manufacturing-Divides.pdf

Andreou, A., Venkatadri, G., Goga, O., Gummadi, K., Loiseau, P. & Mislove, A. (2018). 'Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations'. *Network and Distributed System Security Symposium*, February 2018. https://hal.archives-ouvertes.fr/hal-01955309/

Annenberg Public Policy Center of the University of Pennsylvania. (2020). 'Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online'. https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/07/Freedom_and_Accountability_TWG_Final_Report.pdf

Apelblat, M. (2020). 'UN marks World Press Freedom Day amid attacks against journalists'. *The Brussels Times*, 05 May 2020. https://www.brusselstimes.com/all-news/world-all-news/109683/un-marks-world-press-freedom-day-amid-attacks-against-journalists/

Argentina Political Party Financing Law. (2019). 'Ley De Financiamiento De Los Partidos Políticos Ley 27504'. https://www.boletinoficial.gob.ar/detalleAviso/primera/208603/20190531

Argentinian Government. (2018a). 'Argentina Proyecto de ley Creación de Comisión de Verificación de Noticias Falsas – 2018'. https://www.hcdn.gob.ar/proyectos/textoCompleto.jsp?exp=5228-D-2018&tipo=LEY

Aro, J. (2016). 'The cyberspace war: propaganda and trolling as warfare tools'. *European view* 15.1, 1 June 2016: 121-132. https://journals.sagepub.com/doi/full/10.1007/s12290-016-0395-5

Argentinian Government. (2018b). 'Bill to create a Commission for the Verification of Fake News' (unofficial translation). https://observatoriolegislativocele.com/argentina-proyecto-de-ley-creacion-de-comision-de-verificacion-de-noticias-falsas-2018/

Arthur, R. (2019). 'We Analyzed More Than 1 Million Comments on 4chan. Hate Speech There Has Spiked by 40% Since 2015'. *Vice*, 10 July 2019. https://www.vice.com/en_us/article/d3nbzy/we-analyzed-more-than-1-million-comments-on-4chan-hate-speech-there-has-spiked-by-40-since-2015

Article 19. (2018a). 'YouTube Community Guidelines: Legal Analysis', September 2018. https://www.article19.org/wp-content/uploads/2018/09/YouTube-Community-Guidelines-August-2018.pdf

Article 19. (2018b). 'Facebook Community Guidelines: Legal Analysis', June 2018. https://www.article19.org/wp-content/uploads/2018/07/Facebook-Community-Standards-August-2018-1-1.pdf

Article 19. (2020a). 'Ensuring the Public's Right to Know in the COVID-19 Pandemic'. May 2020. https://www.article19.org/wp-content/uploads/2020/05/Ensuring-the-Publics-Right-to-Know-in-the-Covid-19-Pandemic_Final.pdf

Article 19. (2020b). 'We share the concern of many at reports today of the intention of the US President to sign an executive order that will seek to curtail the free speech protections….' *Twitter*, 28 May 2020. https://twitter.com/article19org/status/1266017734925656065?s=20

ASEAN. (2020). 'Statement of the Special ASEAN-China Foreign Ministers' Meeting on the Coronavirus Disease 2019 (COVID-19)'. 20 February 2020. https://asean.org/storage/2020/02/ASEAN-China-SFMM-Statement-on-COVID-19-20-Feb-2020-Final.pdf

ASEAN AMRI. (2018). 'The ASEAN Ministers Responsible for Information (AMRI) Framework and Joint Declaration to Minimise the Harmful Effects of Fake News'. 10 May 2018. https://asean.org/storage/2012/05/Annex-5-Framework-Declr-Fake-News.pdf

Aspray, W. & Cortada, J.W. (2019). 'From Urban Legends to Political Fact-Checking'. Springer.

Associated Press. (2018). '3 Myanmar journalists in court over story gov't calls false'. 17 October 2018. https://apnews.com/e7d951b784ac48208000d2609435dbab?utm_source=Pew+Research+Center&utm_campaign=c32d4eb996-EMAIL_CAMPAIGN_2018_10_18_01_28&utm_medium=email&utm_term=0_3e953b9b70-c32d4eb996-400451153

Associated Press. (2019). 'Bahrain charges lawyer of sharing "fake news" for his Tweets'. *The Washington Post*, 15 May 2019. https://www.washingtonpost.com/world/middle_east/bahrain-charges-lawyer-of-sharing-fake-news-for-his-tweets/2019/05/15/524e4a04-7727-11e9-a7bf-c8a43b84ee31_story.html

Avaaz. (2019). 'Far Right Networks of Deception'. 22 May 2019. https://avaazimages.avaaz.org/Avaaz%20Report%20Network%20Deception%2020190522.pdf?slideshow

Avaaz. (2020). 'Facebook's Algorithm: A Major Threat to Public Health'. 19 August 2020. https://secure.avaaz.org/campaign/en/facebook_threat_health/

Axelrod, T. (2020). 'Facebook donating $2M to local newsrooms, fact-checkers covering coronavirus'. *The Hill*, 17 March 2020. https://thehill.com/policy/technology/488065-facebook-donating-1-million-to-local-newsrooms-covering-coronavirus-and

Ayyub, R. (2018). "In India, journalists face slut-shaming and rape threats." *New York Times*, 22 May 2018. https://www.nytimes.com/2018/05/22/opinion/india-journalists-slut-shaming-rape.html

Babakar, M. & Moy, W. (2016) 'The State of Automated Factchecking'. *Full Fact report*, 17 August 2016. https://fullfact.org/blog/2016/aug/automated-factchecking/

Babu, A., Lui, A., & Zhang, J. (2017). 'New Updates to Reduce Clickbait Headlines', *Facebook Newsroom*, 17 May 2017. https://newsroom.fb.com/news/2017/05/news-feed-fyi-new-updates-to-reduce-clickbait-headlines/

Ball, J. (2018). 'Post-truth: How bullshit conquered the world'. *Biteback Publishing*, 2017.

Bälz, K. & Mujally, H. (2019). 'Egypt: The New Egyptian Anti-Cybercrime Law Regulates Legal Responsibility For Web Pages And Their Content'. Mondaq, 1 July 2019. http://www.mondaq.com/x/—820028/Security/The+New+Egyptian+AntiCybercrime+Law+Regulates+Legal+Responsibility+for+Web+Pages+and+Their+Content

Bangkok Post. (2019). 'Anti-fake news centre hails first arrest'. 14 November 2019. https://www.bangkFokpost.com/thailand/general/1793649/anti-fake-news-centre-hails-first-arrest

Bangladesh Digital Security Act. (2018). Act No 46 of the Year 2018. 08 October 2018. https://www.cirt.gov.bd/wp-content/uploads/2018/12/Digital-Security-Act-2018-English-version.pdf

Barnathan, J. (2020). 'A Global Crisis Like COVID-19 Calls for a Global Response. Here's Ours.' *ICFJ*, 20 March 2020. https://www.icfj.org/news/global-crisis-covid-19-calls-global-response-heres-ours

Barnett, S. (2016). 'How our mainstream media failed democracy'. In EU Referendum Analysis 2016: Media, Voters and the Campaign. Jackson, D., Thorsen, E., Wring, D., Loughborough University Center for the Study of Journalism, Culture and Community. http://www.referendumanalysis.eu/

Baron Cohen, S. (2019). 'Never Is Now 2019. ADL International Leadership Award Presented to Sacha Baron Cohen'. Recording of his speech, *YouTube*, 21 November 2019. https://www.youtube.com/watch?v=ymaWq5yZIYM&feature=youtu.be

Bartlett, J., Reffin, J., Rumball, N., & Williamson, S. (2014). 'Anti-social media'. *Technical report, Demos*, February 2014. https://www.demos.co.uk/files/DEMOS_Anti-social_Media.pdf

BBC. (2018a). 'Bahrain activist jailed for five years over Twitter comments'. 21 February 2018. https://www.bbc.co.uk/news/world-middle-east-43140519

BBC. (2018b). 'Egypt sentences activist for 'spreading fake news''. 29 September 2018. https://www.bbc.co.uk/news/world-middle-east-45691770

BBC. (2018c). 'Beyond Fake News. BBC launches huge new international anti-disinformation initiative'. 09 November 2018. https://www.bbc.co.uk/mediacentre/latestnews/2018/beyond-fake-news

BBC. (2019a). 'Ivorian MP Alain Lobognan jailed for 'fake news' tweet'. 30 January, 2019. https://www.bbc.co.uk/news/world-africa-47057509

BBC. (2019b). 'Russia internet: Law introducing new controls comes into force'. 01 November 2019. https://www.bbc.co.uk/news/world-europe-50259597

BBC. (2020a). 'BBC Young Reporter and British Council team up to fight 'fake news''. 16 January 2020. https://www.bbc.co.uk/mediacentre/latestnews/2020/young-reporter-fake-news

BBC. (2020b). 'Coronavirus: World leaders' posts deleted over fake news'. 31 March 2020. https://www.bbc.co.uk/news/technology-52106321

BBC. (2020c). 'Twitter hides Trump tweet for 'glorifying violence'. 29 May 2020. https://www.bbc.co.uk/news/technology-52846679

BBC. (2020d). 'Facebook and Twitter restrict Trump accounts over 'harmful' virus claim'. 06 August 2020. https://www.bbc.co.uk/news/election-us-2020-53673797

Beavers, O. (2019). ' Experts are studying mannerisms of 2020 candidates to help offset threat of 'deepfake' videos'. *The Hill*, 09 May 2019. https://thehill.com/policy/cybersecurity/443018-experts-are-studying-mannerisms-of-2020-candidates-to-help-offset-threat

Beckett, C. (2016). 'Deliberation, distortion and dystopia: the news media and the referendum'. *In EU Referendum Analysis 2016: Media, Voters and the Campaign. Jackson, D, Thorsen, E, Wring, D. Loughborough University Center for the Study of Journalism, Culture and Community*. http://www.referendumanalysis.eu/

Bell, E. & Owen, T. (2017). 'The Platform Press: How Silicon Valley Reengineered Journalism'. *Tow Center for Digital Journalism*, 29 March 2017. https://www.cjr.org/tow_center_reports/platform-press-how-silicon-valley-reengineered-journalism.php

Bengani, P. (2020). 'As election looms, a network of mysterious 'pink slime' local news outlets nearly triples in size'. *CJR*, 4 August 2020. https://www.cjr.org/analysis/as-election-looms-a-network-of-mysterious-pink-slime-local-news-outlets-nearly-triples-in-size.php

Benin Digital Code, (2017). Assemblee Nationale. Loi n° 2017-20 portant code du numérique en République du Bénin. 13 June 2017. https://www.afapdp.org/wp-content/uploads/2018/06/Benin-Loi-2017-20-Portant-code-du-numerique-en-Republique-du-Benin.pdf

Benkler, Y., Faris, B., Roberts, H. (2018). 'Network Propaganda. Manipulation, Disinformation, and Radicalization in American Politics'. *Oxford University Press*.

Benton, J. (2020). 'Is this video "missing context," "transformed," or "edited"? This effort wants to standardize how we categorize visual misinformation'. *NiemanLab*, 16 January 2020. https://www.niemanlab.org/2020/01/is-this-video-missing-context-transformed-or-edited-this-effort-wants-to-standardize-how-we-categorize-visual-misinformation/

Beo Da Costa, A. (2019). 'Indonesia lifts social media curbs targeting hoaxes during unrest'. *Reuters*, 25 May 2019. https://www.reuters.com/article/us-indonesia-election-social-media/indonesia-lifts-social-media-curbs-targeting-hoaxes-during-unrest-idUSKCN1SV06J

Berger, G. (2019). 'Whither MIL: Thoughts for the Road Ahead'. Chapter in 'Understanding Media and Information Literacy (MIL) in the Digital Age. A Question of Democracy'. (ed). Ulla Carlsson, Goteborg: Nordicom. https://jmg.gu.se/digitalAssets/1742/1742676_understanding-media-pdf-original.pdf

Bhattacharjee, S. & Dotto, C. (n.d.) 'First Draft case study: Understanding the impact of polio vaccine disinformation in Pakistan'. *First Draft*. https://firstdraftnews.org/long-form-article/first-draft-case-study-understanding-the-impact-of-polio-vaccine-disinformation-in-pakistan/

Bich Ngoc, N. (2019). 'Vietnam's New Cybersecurity Law 2018'. *Vietnam Business Law*. 30 July 2018. https://vietnam-business-law.info/blog/2018/7/30/vietnams-new-cybersecurity-law

Bickert, M. (2018). 'Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process', *Facebook Newsroom*, 24 April 2018. https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/

Bickert, M. (2019). 'Updating the Values That Inform Our Community Standards', *Facebook Newsroom*, 12 September 2019. https://about.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards/

Bigot, L. (2019). Fact-Checking vs. fake news. Vérifier pour mieux informer. *INA*, 18 October 2019. https://presse.ina.fr/fact-checking-vs-fake-news/

Billing, L. (2020). 'Duterte's troll armies drown out Covid-19 dissent in the Philippines'. *Coda*, 21 July 2020. https://www.codastory.com/disinformation/philippines-troll-armies/

Bilton, R. (2016). 'Electionland, a joint project between ProPublica and six other orgs, will cover Election Day voting issues'. *NiemanLab*, 08 September 2016. https://www.niemanlab.org/2016/09/electionland-a-joint-project-between-propublica-and-six-other-orgs-will-create-a-virtual-newsroom-to-cover-election-day-voting-issues/

Binder, M. (2019). 'Facebook ad scam tricks users with images and video of Kickstarter products'. *Mashable UK*, 15 November 2019. https://mashable.com/article/facebook-scam-crowdfunding-ads.amp/?europe=true

Birks, J. (2019). '*Fact-Checking Journalism and Political Argumentation. A British Perspective*'. Palgrave / MacMillan

Board, J. (2019). 'Inside Indonesia's 'fake news' war room, fighting political hoaxes in election season'. *CAN*, 13 April 2019. https://www.channelnewsasia.com/news/asia/indonesia-election-fake-news-war-room-fighting-political-hoaxes-11439398

Boothroyd-Rojas, R. (2017). 'Venezuela's Constituent Assembly Passes Anti-Hate Crime Law'. *Venezuelanalysis*, 09 November 2017. https://venezuelanalysis.com/news/13492

Bosworth, A. (2020). 'Lord of the Rings, 2020 and Stuffed Oreos: Read the Andrew Bosworth Memo'. *New York Times*, 7 January 2020. https://www.nytimes.com/2020/01/07/technology/facebook-andrew-bosworth-memo.html

Bossetta, M. (2018). 'The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election'. *Journalism & Mass Communication Quarterly*, 28 March 2018. https://doi.org/10.1177/1077699018763307

Bozdag, E., & Van den Hoven, M. J. (2015). 'Breaking the filter bubble: Democracy and design'. *Ethics and Information Technology. 17.4, 249-265*, 18 December 2015. https://repository.tudelft.nl/islandora/object/uuid%3Ae24751ba-b94b-4856-b4a9-2f5f4f25ff14

Bradshaw, S. & Howard, P.N. (2018). 'Challenging Truth and Trust: a Global Inventory of Organized Social Media Manipulation'. *Working Paper 2018.1. Oxford, UK: Project on Computational Propaganda, 26 pp*, 20 July 2018. https://comprop.oii.ox.ac.uk/research/cybertroops2018/

Bradshaw, S. & Howard, P.N. (2019). 'The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation'. *Oxford, UK: Project on Computational Propaganda, 26 September 2019*. https://comprop.oii.ox.ac.uk/research/cybertroops2019/

Brazil Superior Electoral Court. (2018). 'Fake News: TSE lança página para esclarecer eleitores', Press Release, 11 October 2018. http://www.tse.jus.br/imprensa/noticias-tse/2018/Outubro/fake-news-tse-lanca-pagina-para-esclarecer-eleitores-sobre-a-verdade

Brennen, J. S., Simon, F., Howard, P. N. & Nielsen, R. K. (2020). 'Types, sources, and claims of COVID-19 misinformation'. *Reuters*, 07 April 2020. https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation

Broadband Commission. (2013). Technology, broadband and education: Advancing the education for all agenda. *UNESCO Publication*. https://unesdoc.unesco.org/ark:/48223/pf0000219687.locale=en

Brooking, E.T., Kann, A. & Rizzuto, M. (2020). 'Dichotomies of Disinformation'. *DFRLab*. https://github.com/DFRLab/Dichotomies-of-Disinformation/blob/master/README.md

Buchanan, K. (2019). 'Initiatives to Counter Fake News: Malaysia'. *Library of Congress*, April 2019. https://www.loc.gov/law/help/fake-news/malaysia.php

Buchanan, T. & Benson, V. (2019). 'Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of "Fake News"?'. *Social Media + Society*, 17 December 2019. https://doi.org/10.1177/2056305119888654

Buckmaster, L. & Wils, T. (2019). 'Responding to Fake News. *Australia Parliamentary Library Briefing*'. https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/BriefingBook46p/FakeNews

Bukhari, P. (2019). 'Srinagar, India | AFP | Monday 12/23/2019 – 15:28 UTC+8 | 663 words'. *The Independent News Singapore*, 23 December 2019. http://theindependent.sg/srinagar-india-afp-monday-12-23-2019-1528-utc8-663-words/

Buning, M.D.C. et al. (2018). 'A multi-dimensional approach to disinformation'. Report of the independent High Level Group on Fake News and Online Disinformation'. *European Commission*, 12 March 2018. https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

Burgess, M. (2018). 'To fight fake news on WhatsApp, India is turning off the internet'. *Wired*, 18 October 2018. https://www.wired.co.uk/article/whatsapp-web-internet-shutdown-india-turn-off

Burgos, P. (2019). 'What 100,000 WhatsApp Messages Reveal about Misinformation in Brazil', *First Draft*, 27 June 2019. https://firstdraftnews.org/latest/what-100000-whatsapp-messages-reveal-about-misinformation-in-brazil/

Burki, T. (2019). 'Vaccine misinformation and social media'. *The Lancet*, 01 October 2019. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(19)30136-0/fulltext

Burkina Faso Constitutional Council. (2019). 'Décision n° 2019- 013/CC sur le contrôle de constitutionnalité de la loi n° 044-2019/AN du 21 juin 2019 portant modification de la loi n° 025-2018/AN du 31 mai 2018 portant Code pénal par autosaisine'. https://www.conseil-constitutionnel.gov.bf/fileadmin/user_upload/decision_13__code_penal.pdf

Cadwalladr, C. (2017a). Revealed: Tory 'dark' ads targeted voters' Facebook feeds in Welsh marginal seat. *The Observer*. 27 May 2017. https://www.theguardian.com/politics/2017/may/27/conservativesfacebook-dark-ads-data-protection-election

Cadwalladr, C. (2017b). 'The great British Brexit robbery: how our democracy was hijacked'. *The Guardian*, 07 May 2017. https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

Cadwalladr, C. (2018). ''Plucky little panel' that found the truth about fake news Facebook and Brexit'. *The Guardian*, 28 Jul 2018. https://www.theguardian.com/politics/2018/jul/28/dcms-committee-reportfinds-truth-fake-news-facebook-brexit

Cameroon Cyber Security and Cyber Criminality Law. (2010). 'Loi n°2010/012 du 21 Decembre 2010 relative a la Cybersecurite et la Cybercriminalite au Cameroun'. https://www.unodc.org/res/cld/document/cmr/2010/loi_sur_la_cybersecurite_et_la_cybercriminalite_html/Loi_2010-012_cybersecurite_cybercriminalite.pdf

Cameroon Penal Code (1967 Revised). https://www.wipo.int/edocs/lexdocs/laws/en/cm/cm014en.pdf

Canada House of Commons Standing Committee on Access to Information, Privacy and Ethics. (2018). 'Democracy under Threat: Risks and Solutions in the Era of Disinformation and Data Monopoly'. *Ottowa*: *House of Commons*. https://www.ourcommons.ca/DocumentViewer/en/42-1/ETHI/report-17

Canada House of Commons Standing Committee on Access to Information, Privacy and Ethics. (2019). *International Grand Committee on Big Data, Privacy and Democracy. Report and Government Response*. https://www.ourcommons.ca/Committees/en/ETHI/StudyActivity?studyActivityId=10554743

Canadian Government. (2018). 'Elections Modernization Act'. https://laws-lois.justice.gc.ca/eng/annualstatutes/2018_31/page-1.html

Canadian Government. (2019a). *Response to Canadian Parliamentary Standing Committee on Access to Information, Privacy and Ethics Report. Democracy under Threat*: *Risks and Solutions in the Era of Disinformation and Data Monopoly*. https://www.ourcommons.ca/DocumentViewer/en/42-1/ETHI/report-17/response-8512-421-502

Canadian Government. (2019b). *Online Disinformation*. https://www.canada.ca/en/canadian-heritage/services/online-disinformation.html

Canadian Government. (2019c). 'Backgrounder – Helping Citizens Critically Assess and Become Resilient Against Harmful Online Disinformation'. *Canadian Heritage*, 21 August 2019. https://www.canada.ca/en/canadian-heritage/news/2019/07/backgrounder--helping-citizens-critically-assess-and-become-resilient-against-harmful-online-disinformation.html

Canadian Heritage. (2020). 'COVID-19: The Government of Canada is taking action to support the publishing and news sectors'. https://www.canada.ca/en/canadian-heritage/news/2020/03/covid-19-the-government-of-canada-is-taking-action-to-support-the-publishing-and-news-sectors.html

Carey, J. M., Chi, V., Flynn, D. J., Nyhan, B. & Zeitoff, T. (2020). 'The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil'. Science Advances, January 2020. DOI: 10.1126/sciadv.aaw7449

Carmi, E., Yates, S.J., Lockley, E. & Pawluczuk, A. (2020). 'Data citizenship: rethinking data literacy in the age of disinformation, misinformation, and malinformation'. *Internet Policy Review*, 28 May 2020. https://policyreview.info/articles/analysis/data-citizenship-rethinking-data-literacy-age-disinformation-misinformation-and

Carmichael, F. & Gragnani, J. (2019). 'YouTube advertises big brands alongside fake cancer cure videos'. *BBC*, 13 September 2019. https://www.bbc.co.uk/news/blogs-trending-49483681

Carmichael, F. & Hussain, A. (2019). 'Pro-Indian 'fake websites targeted decision makers in Europe'. *BBC*, 16 December 2019. https://www.bbc.co.uk/news/world-asia-india-50749764

Cassini, S. (2019). 'Christophe Castaner et la Pitié-Salpêtrière : premier désaveu pour la loi sur les « infox »'. *Le Monde*, 22 May 2019. https://www.lemonde.fr/politique/article/2019/05/22/premier-desaveu-pour-la-loi-sur-les-infox_5465717_823448.html

Centre for Data Ethics and Innovation. (2019). 'Deepfakes and Audio-visual Disinformation'. September 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/831179/Snapshot_Paper_-_Deepfakes_and_Audiovisual_Disinformation.pdf

Center for Human Rights, United Nations. (1994). 'Handbook on the Legal, Technical, and Human Rights Aspects of Elections'. https://eos.cartercenter.org/uploads/document_file/path/8/training2enTCCoptimized.pdf

Chakraborty, A., Sarkar, R., Mrigen, A., & Ganguly, N. (2017). 'Tabloids in the Era of Social Media? Understanding the Production and Consumption of Clickbaits in Twitter'. SSRN, 10 July 2017. http://dx.doi.org/10.2139/ssrn.3034591

Chaturvedi, S. (2016). 'I am a troll: Inside the secret world of the BJP's digital army'. *Juggernaut Books*, 2016.

Chaturvedi, S. (2020). 'Govt launches chatbot on WhatsApp to create awareness about coronavirus, curb misinformation'. *The Economic Times*, 22 March 2020. https://economictimes.indiatimes.com/tech/internet/govt-launches-chatbot-on-whatsapp-to-create-awareness-about-coronavirus-curb-misinformation/articleshow/74750648.cms

Chavoshi, N., Hamooni, H., & Mueen, A. (2017). 'Temporal Patterns in Bot Activities'. *ACM*, April, 2017. http://doi.org/10.1145/3041021.3051114

Checknews (2019). '*Combien a rapporté à Libé son partenariat de factchecking avec Facebook en 2018*?'. 30 January 2019. https://www.liberation.fr/checknews/2019/01/30/combien-a-rapporte-a-libe-son-partenariat-de-factchecking-avec-facebook-en-2018_1706160

Chile Senate. (2019). "Fake news": autoridades nacionales y regionales podrían tener nueva causal para la cesación de cargos. Press Release, 7 February 2019. https://www.senado.cl/fake-news-autoridades-nacionales-y-regionales-podrian-tener-nueva/senado/2019-01-29/122725.html

Chitranukroh, A. (2017). 'The new Computer Crimes Act and concerns over online freedom'. *Bangkok Post*, 20 January 2017. https://www.bangkokpost.com/business/1183561/the-new-computer-crimes-act-and-

Christopher, N. (2020). 'We've Just Seen the First Use of Deepfakes in an Indian Election Campaign'. *Vice*, 18 February 2020. https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp

Chung Seck, Y. & Son Dang, T. (2019). 'Vietnam National Assembly Passes the Law on Cybersecurity'. *Global Compliance News*, 02 July 2018. https://globalcompliancenews.com/vietnam-law-cybersecurity-20180702/

Clegg, N. (2019). 'Facebook, Elections and Political Speech'. *Facebook*, 24 September 2019. https://about.fb.com/news/2019/09/elections-and-political-speech/

Clegg, N. (2020). 'Welcoming the Oversight Board'. *Facebook*, 06 May 2020. https://about.fb.com/news/2020/05/welcoming-the-oversight-board/

Cohen, N. (2020). 'How Wikipedia Prevents the Spread of Coronavirus Misinformation'. *Wired*, 15 March 2020. https://www.wired.com/story/how-wikipedia-prevents-spread-coronavirus-misinformation/

Collingridge, D. (1980). 'The social control of technology'. St Martin, New York

Conseil Constitutionnel. (2018). 'Décision n° 2018-773 DC du 20 décembre 2018'. https://www.conseil-constitutionnel.fr/decision/2018/2018773DC.htm

Cook, J., Ecker, U.K.H & Lewandowsky, S. (2014). 'Misinformation and How to Correct It.' *ResearchGate*, July 2014. https://www.researchgate.net/publication/265643150_Misinformation_and_How_to_Correct_It

Cook, J., van der Linden, S., Lewandowsky, S. & Ecker, U. (2020). 'Coronavirus, 'Plandemic' and the seven traits of conspiratorial thinking'. *The Conversation*, 15 May 2020. https://theconversation.com/coronavirus-plandemic-and-the-seven-traits-of-conspiratorial-thinking-138483

Corbu, N., Oprea, D-A., Negrea-Busuioc, E. & Radu, L. (2020). ''They can't fool me, but they can fool the others!' Third person effect and fake news detection'. *European Journal of Communication*, 17 February 2020. https://journals.sagepub.com/doi/full/10.1177/0267323120903686?journalCode=ejca

Corcoran, C. Crowley, B.J., Davis, R. (2019) 'Disinformation Threat Watch. The Disinformation Landscape in East Asia and Implications for US Policy'. *Harvard Kennedy School, Student Report*, May 2019. https://www.belfercenter.org/sites/default/files/2019-06/PAE/DisinfoWatch%20-%202.pdf

Cordey, S. (2019). 'Cyber Influence Operations: An Overview and Comparative Analysis'. *Zurich: Center for Security Studies* (*CSS*). https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Cyber-Reports-2019-10-CyberInfluence.pdf

Côte d'Ivoire Penal Code 1981, revised. Article 97, 2017. http://www.caidp.ci/uploads/01981c9a7d883c4321811e8725ca4c2c.pdf

Côte d'Ivoire Penal Code 1981, revised. Article 173. http://www.gouv.ci/doc/accords/1512502410CODE-PENAL.pdf

CPJ. (2019a). 'Cameroonian journalist detained on criminal defamation and false news charges'. 20 June 2019. https://cpj.org/2019/06/critical-cameroonian-journalist-detained-on-crimin.php

CPJ. (2019b). 'Journalist Ignace Sossou convicted of false news in Benin'. 23 August 2019. https://cpj.org/2019/08/journalist-ignace-sossou-convicted-of-false-news-i.php

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A. & Tesconi, M. (2016). 'DNA-inspired online behavioral modeling and its application to spambot detection'. *IEEE Intelligent Systems*, 31(5), 30 January 2016. https://www.researchgate.net/publication/298902745_DNA-Inspired_Online_Behavioral_Modeling_and_Its_Application_to_Spambot_Detection

Damiano Ricci, A. (2018). 'French opposition parties are taking Macron's anti-misinformation law to court'. *Poynter*, 04 December 2018. https://www.poynter.org/fact-checking/2018/french-opposition-parties-are-taking-macrons-anti-misinformation-law-to-court/

Danish Ministry of Foreign Affairs. (2018). 'Strengthened safeguards against foreign influence on Danish elections and democracy'. Press Release, 07 September 2018. https://um.dk/en/news/newsdisplaypage/?newsid=1df5adbb-d1df-402b-b9ac-57fd4485ffa4

Dara, V. (2019). 'Government to launch TV show against fake news'. *The Phnom Penh Post*, 24 January 2019. https://www.phnompenhpost.com/national/government-launch-tv-show-against-fake-news

Darmanin, J. (2019). 'FCEU Newsletter #7 - Good news and bad news after the election week-end'. *Fact Check EU*, 27 May 2019. https://factcheckeu.info/en/article/fceu-newsletter-7-good-news-and-bad-news-after-election-week-end

Daudin, G. (2019). '52% des viols commis à Paris en 2014 l'ont été par des étrangers. L'étude ne dit pas cela'. *AFP Factuel*, 25 November 2019. https://factuel.afp.com/52-des-viols-commis-paris-en-2014-lont-ete-par-des-etrangers-letude-ne-dit-pas-cela

Davis, A. (2019). Political Communication. A New Introduction for Crisis Times. Polity.

Dawn. (2018). 'Govt launches 'Fake News Buster' account to expose false reports'. 01 October 2018. https://www.dawn.com/news/1436167

Dean, M. (2017). 'Snopes and the Search for Facts in a Post-Fact World'. *Wired*, 20 September 2017. https://www.wired.com/story/snopes-and-the-search-for-facts-in-a-post-fact-world/

Decker, B. (2019). 'Adversarial Narratives: A New Model for Disinformation'. *GDI*, August 2019. https://disinformationindex.org/wp-content/uploads/2019/08/GDI_Adverserial-Narratives_Report_V6.pdf

DCMS HC 363 (2018). 'Oral evidence : Fake News – 8 February 2018 (George Washington University, Washington DC), HC 363'. 08 February 2018. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/fake-news/oral/78195.html

De Croo, A. (2018). 'Fact checking fonds in de steigers in strijd tegen fake news'. Press Release, 08 October 2018. https://alexanderdecroo.be/fact-checking-fonds-steigers-strijd-fake-news/

Derakhshan, D. (2019). 'Disinfo Wars. A taxonomy of information warfare'. *Medium*, 9 May 2019. https://medium.com/@h0d3r/disinfo-wars-7f1cf2685e13

Dewey, C. (2016). 'Facebook Fake-News Writer: 'I Think Donald Trump is in the White House because of Me.'' *The Washington Post*, November 17, 2016. https://www.washingtonpost.com/news/the-intersect/wp/2016/11/17/facebook-fake-news-writer-i-think-donald-trump-is-in-the-white-house-because-of-me/

DiResta, R. (2018). 'Free Speech Is Not the Same As Free Reach'. *Wired*, 30 August 2018. https://www.wired.com/story/free-speech-is-not-the-same-as-free-reach/

Dobber, T., Ó Fathaigh, R. & Zuiderveen Borgesius, F. (2019). 'The regulation of online political micro-targeting in Europe'. *Internet Policy Review*, 31 December 2019. https://policyreview.info/articles/analysis/regulation-online-political-micro-targeting-europe

Dorroh, J. (2020). Beyond fact-checking: fighting the onslaught of COVID-19 disinformation. Ijnet international journalists' network, 2 July 2020. https://ijnet.org/en/story/beyond-fact-checking-fighting-onslaught-covid-19-disinformation

Dorsey, J. (2019). 'We've made the decision to stop all political advertising on Twitter globally. We believe political message reach should be earned, not bought. Why? A few reasons…' *Twitter*, 30 October 2019. https://twitter.com/jack/status/1189634360472829952?s=20

Douek, E. (2020). 'COVID-19 and Social Media Content Moderation'. *Lawfare*, 25 March 2020, https://www.lawfareblog.com/covid-19-and-social-media-content-moderation

Doyle, A. (2016). 'Facebook Says Will Learn From Mistake Over Vietnam Photo'. *Hindustan Times*, 13 September 2016. http://www.hindustantimes.com/world-news/facebook-says-will-learn-from-mistake-over-vietnam-photo/story-kwmb3iX6lKgmwaIGZeKlyN.html

Drissa, D. (2019). 'Offensive générale contre les fausses informations sur les réseaux sociaux'. *7info*, 12 July 2019. https://www.7info.ci/offensive-generale-contre-les-fausses-informations-sur-les-reseaux-sociaux/

Drobnic Holan, A. (2018). 'Politifact: The Principles of the Truth-O-Meter: PolitiFact's methodology for independent fact-checking'. *Politifact*, 12 February 2018. https://www.politifact.com/truth-o-meter/article/2018/feb/12/principles-truth-o-meter-politifacts-methodology-i/

Duarte, N. & Llansó, E. (2017). 'Mixed Messages: The Limits of Automated Social Media Content Analysis'. *Center for Democracy and Technology*, 28 November 2017. https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/

Dubois, E., & Blank, G. (2018). 'The echo chamber is overstated: the moderating effect of political interest and diverse media'. Information, Communication & Society, 29 January 2018. 21(5), 729-745. https://doi.org/10.1080/1369118X.2018.1428656

Dufour, N. & Gully, A. (2019). 'Contributing Data to Deepfake Detection Research'. *Google Blog*, 24 September 2019. https://ai.googleblog.com/2019/09/contributing-data-to-deepfake-detection.html

Durach, F. (2020). 'Who's afraid of fake news? New evidence from Romania'. *YouCheck*, 24 February 2020. http://project-youcheck.com/whos-afraid-of-fake-news-new-evidence-from-romania/

Dutch Government (2019a). Fake news campaign starts today. Press Release, 11 March 2019. https://www.rijksoverheid.nl/onderwerpen/desinformatie-nepnieuws/nieuws/2019/03/11/campagne-nepnieuws-vandaag-van-start

Dutch Government (2019b). Kabinet zet in op transparantie in strategie tegen desinformatie. Press Release, 18 October 2019). https://www.rijksoverheid.nl/onderwerpen/desinformatie-nepnieuws/nieuws/2019/10/18/kabinet-zet-in-op-transparantie-in-strategie-tegen-desinformatie

Dutch Government (2019c). Actielijnen tegengaan desinformatie. https://www.rijksoverheid.nl/onderwerpen/desinformatie-nepnieuws/documenten/kamerstukken/2019/10/18/actielijnen-tegengaan-desinformatie

Dwoskin, E., Whalen, J. & Cabato, R. (2019). 'Content moderators at YouTube Facebook and Twitter see the worst of the web – and suffer silently. *The Washington Post*, 25 July 2019. https://www.washingtonpost.com/technology/2019/07/25/social-media-companies-are-outsourcing-their-dirty-work-philippines-generation-workers-is-paying-price/

EC Communication on Tackling Illegal Content Online. (COM (2017) 555 final) https://webcache.googleusercontent.com/search?q=cache:ZJcX6Dn0020J:https://ec.europa.eu/transparency/regdoc/rep/1/2017/EN/COM-2017-555-F1-EN-MAIN-PART-1.PDF+&cd=1&hl=en&ct=clnk&gl=uk&client=firefox-b-e

ECI. (2019). 'Report of the Committee on Section 126 of the Representation of the People Act, 1951 Submitted to The Commission'. 10 January 2019. https://pib.gov.in/newsite/PrintRelease.aspx?relid=187412

Ecker, U. K. H., Lewandowsky, S, and Tang, D.T.W. (2010). "Explicit warnings reduce but do not eliminate the continued influence of misinformation". *Memory & Cognition*. Vol.38, No. 8, pp. 1087-1100 https://doi.org/10.3758/MC.38.8.1087

Ecker, U. K. H., O'Reilly, Z., Reid, J. S. & Chang, E. P. (2019). 'The effectiveness of short-format refutational fact-checks'. *British Journal of Psychology*, 02 March 2019. https://onlinelibrary.wiley.com/doi/full/10.1111/bjop.12383

Eco, U. (2014). 'From the Tree to the Labyrinth'. *Harvard University Press*.

Effron, D. A. & Raj, M. (2019). 'Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share'. *Psychological Science*, *31*(1), 75–87. https://doi.org/10.1177/0956797619887896

EFJ (2018). 'Belarus: more media censorship and control with new amendments of Media Law'. 24 June 2018. https://europeanjournalists.org/blog/2018/06/24/belarus-more-media-censorship-and-control-with-new-amendments-of-the-media-law/

Eisenstat, Y. (2019). 'I worked on political ads at Facebook. They profit by manipulating us.' The Washington Post, 04 November 2019. https://www.washingtonpost.com/outlook/2019/11/04/i-worked-political-ads-facebook-they-profit-by-manipulating-us/

El Khoury, R. (2020). 'WhatsApp has rolled out the restriction of single chat forwards for viral messages'. *Android Police*, 13 April 2020. https://www.androidpolice.com/2020/04/13/whatsapp-will-limit-viral-message-forwards-to-one-chat-at-a-time-to-combat-coronavirus-misinformation/

Ellis, S. (1989). 'Tuning in to pavement radio.' *African Affairs* 88.352, July 1989 pp321-330 https://pdfs.semanticscholar.org/d07a/60d84e60248e9523c213c7103cd27ce84f8e.pdf

Ellis-Petersen, H. (2019). 'Social media shut down in Sri Lanka in bid to stem misinformation'. *The Guardian*, *21* April 2019. https://www.theguardian.com/world/2019/apr/21/social-media-shut-down-in-sri-lanka-in-bid-to-stem-misinformation

Emarketer. (2020). 'US Political Ad Spending to Hit Record High'. 12 February 2020. https://www.emarketer.com/newsroom/index.php/us-political-ad-spending-to-hit-record-high/

Embury-Dennis, T. (2020). 'Coronavirus: New York sees spike in disinfectant exposure cases following Trump's dangerous treatment musings'. *Independent*, 25 April 2020. https://www.independent.co.uk/news/world/americas/us-politics/coronavirus-trump-treatment-disinfectant-bleach-new-york-a9483786.html

Endeshaw, D. (2020). 'Ethiopia passes law imposing jail terms for internet posts that stir unrest'. *Reuters*, 13 February 2020. https://www.reuters.com/article/us-ethiopia-politics/ethiopia-passes-law-imposing-jail-terms-for-internet-posts-that-stir-unrest-idUSKBN2071PA

Election Monitoring. (2019). 'UK General Election 2019; Digital disruption by the political parties, and the need for new rules'. December, 2019. https://www.isdglobal.org/wp-content/uploads/2019/12/UK-GE-2019-Digital-Disruption-report.pdf

Eljechtimi, A. (2020). 'Morocco makes a dozen arrests over coronavirus fake news'. *Reuters*, 19 March 2020. https://uk.reuters.com/article/uk-health-coronavirus-morocco/morocco-makes-dozen-arrests-over-coronavirus-fake-news-idUKKBN2162EA

Epstein R., Robertson, R. E. (2015). 'The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections'. In *Proceedings of the National Academy of Sciences* (*PNAS*). *112*(*33*):*E4512–E4521*, 18 August 2015. https://www.pnas.org/content/112/33/E4512

Estarque, M. (2020). 'Brazilian projects for media literacy and combating false news find allies outside journalism'. *Knight Center*, 29 January 2020. https://knightcenter.utexas.edu/blog/00-21557-brazilian-projects-media-literacy-and-combating-false-news-find-allies-outside-journal

EU Disinfo Lab (2019a). 'Uncovered: 265 coordinated fake local media outlets serving Indian interests'. 26 November 2019. https://www.disinfo.eu/publications/uncovered-265-coordinated-fake-local-media-outlets-serving-indian-interests

EU Disinfo Lab (2019b). 'An investigation into a pro-Indian influence network'. https://www.disinfo.eu/wp-content/uploads/2019/12/20191213_InfluencingPolicymakers-with-Fake-media-outlets.pdf

EU Disinfo Lab (2019c). 'How you thought you support the animals and you ended up funding white supremacists'. 11 September 2019. https://www.disinfo.eu/publications/suavelos-white-supremacists-funded-through-facebook

EU Disinfo Lab. (2020). 'From health disinformation to copy-pasting Sputnik and RT articles – how an Africa-based network built fake media outlets and clickbait websites for profit'. 8 March 2020. https://www.disinfo.eu/publications/from-health-disinformation-to-copy-pasting-articles-from-sputnik-and-rt-how-an-africa-based-network-built-fake-media-outlets-and-clickbait-websites-for-profit

EU EEAS (2018). 'Questions and Answers about the East Stratcom Task Force'. https://eeas.europa.eu/headquarters/headquarters-homepage/2116/-questions-and-answers-about-the-east-stratcom-task-force_en

EU Foreign Affairs Council. (2018). 'Outcome of The Council Meeting'. 16 April 2018. https://www.consilium.europa.eu/media/33743/st07997-en18.pdf

EU HLEG. (2018). EU High Level Expert Group on Fake News and Online Disinformation. '*A Multi-Dimensional Approach to Disinformation*'. https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

EurActiv. (2016). 'Latvia shuts down Russian 'propaganda' website Sputnik'. 30 March 2016. https://www.euractiv.com/section/global-europe/news/latvia-shuts-down-russias-propaganda-website-sputnik/

EurActiv. (2019). 'France adopts tough law against online hate speech'. 10 July 2019. https://www.euractiv.com/section/politics/news/france-adopts-tough-law-against-online-hate-speech/

Eurobarometer 464. (2018). 'Final results of the Eurobarometer on fake news and online disinformation'. https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-onlinedisinformation

European Commission. (2018a). 'Communication - Tackling online disinformation: a European Approach' (COM(2018) 236 final), 26 April 2018. https://ec.europa.eu/digital-single-market/en/news/communication-tackling-online-disinformation-european-approach

European Commission. (2018b). 'State of the Union 2018: European Commission Proposes Measures for Securing Free and Fair European Elections', Press Release (IP/18/5681), 12 September 2018. http://europa.eu/rapid/press-release_IP-18-5681_en.htm

European Commission. (2018c). 'Code of Practice on Disinformation'. 26 September 2018. https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation

European Commission. (2018d). 'Roadmaps to implement the Code of Practice on disinformation'. 16 October 2018. https://ec.europa.eu/digital-single-market/en/news/roadmaps-implement-code-practice-disinformation

European Commission. (2018e) Joint Communication. '*Action Plan against Disinformation*' *JOIN* (*2018*) *36 final*, 05 December 2018. https://ec.europa.eu/digital-single-market/en/news/action-plan-against-disinformation

European Commission. (2019). Policy 'Tackling Online Disinformation'. 13 September 2019. https://ec.europa.eu/digital-single-market/en/tackling-online-disinformation

European Commission and High Representative. (2018). 'Action Plan against Disinformation'. 05 December 2018. https://ec.europa.eu/commission/sites/beta-political/files/eu-communication-disinformation-euco-05122018_en.pdf

European Parliament. (2020). 'Disinformation: how to recognise and tackle Covid-19 myths'. 30 March 2020. https://www.europarl.europa.eu/news/en/headlines/society/20200326STO75917/disinformation-how-to-recognise-and-tackle-covid-19-myths

EUvsDisinfo. (2020). 'EEAS Special Report: Disinformation on the Coronavirus – short assessment of the information environment'. 19 March 2020. https://euvsdisinfo.eu/eeas-special-report-disinformation-on-the-coronavirus-short-assessment-of-the-information-environment/

Evans, R. (2020). 'How Coronavirus Scammers Hide On Facebook And YouTube'. *Bellingcat*, 19 March 2020. https://www.bellingcat.com/news/rest-of-world/2020/03/19/how-coronavirus-scammers-hide-on-facebook-and-youtube/

Facebook. (2018). 'Facts About Content Control on Facebook'. *Facebook Newsroom*, 28 December 2018. https://about.fb.com/news/2018/12/content-review-facts/

Facebook. (2018b). 'An Independent Assessment of the Human Rights Impact of Facebook in Myanmar'. 05 November 2018. https://about.fb.com/news/2018/11/myanmar-hria/

Facebook. (2019). Ad Library. https://www.facebook.com/ads/library/

Facebook. (2019b). 'Fact-Checking on Facebook: What Publishers should know'. https://www.facebook.com/help/publisher/182222309230722

Facebook. (2019c). Community Standards. https://www.facebook.com/communitystandards/introduction

Facebook. (2019d). Letter from Rebecca Stimson, Head of Public Policy, UK to Department of Digital, Culture, Media and Sport Committee House of Commons on 29 October 2019. https://www.parliament.uk/documents/commons-committees/culture-media-and-sport/191029%20Rebecca%20Stimson%20Facebook%20to%20Chair%20response%20to%2022%20Oct%20letter.pdf

Facebook. (2019e). Oversight Board Charter. https://fbnewsroomus.files.wordpress.com/2019/09/oversight_board_charter.pdf

Facebook. (2020a). 'Working With Industry Partners – Joint industry statement from Facebook, Google, LinkedIn, Microsoft, Reddit, Twitter and YouTube'. 16 March 2020. https://about.fb.com/news/2020/04/coronavirus/#joint-statement

Facebook. (2020b). 'An Update to How We Address Movements and Organizations Tied to Violence'. 19 August 2020. https://about.fb.com/news/2020/08/addressing-movements-and-organizations-tied-to-violence/

Facebook and Instagram. (2019). 'Facebook Report on the Implementation of the Code of Practice for Disinformation – Annual Report'. https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019

FactCheck. (2019). 'Our Funding'. https://www.factcheck.org/our-funding/

Falck, B. (2018). 'Providing More Transparency Around Advertising on Twitter', *Twitter Blog*, 28 June 2018. https://blog.twitter.com/official/en_us/topics/company/2018/Providing-More-Transparency-Around-Advertising-on-Twitter.html

Faridani, S. (2010). 'Opinion space: a scalable tool for browsing online comments'. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (*CHI'10*), *1175-1184*. *ACM*, April 2010. https://dl.acm.org/doi/10.1145/1753326.1753502

Faris, R., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E. & Benkler, Y. (2017). 'Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election'. *Berkman Klein Center Research Publication 2017-6*. https://dash.harvard.edu/handle/1/33759251

Fidler, D. P. (2019). 'Disinformation and Disease: Social Media and the Ebola Epidemic in the Democratic Republic of the Congo'. *Council on Foreign Affairs*, 20 August 2019. https://www.cfr.org/blog/disinformation-and-disease-social-media-and-ebola-epidemic-democratic-republic-congo

Fischer, S. (2020). 'Exclusive: Facebook cracks down on political content disguised as local news'. *Axios*, 11 August 2020. https://www.axios.com/facebook-pages-news-exemption-e66d92ce-2abd-4293-b2ad-16cf223e12f1.html

Fitzgibbon, W. (2020). 'Benin investigative journalist jailed in 'absurd' decision'. *ICIJ*, 19 May 2020. https://www.icij.org/blog/2020/05/benin-investigative-journalist-jailed-in-absurd-decision/

Flaxman, S. R., Goel, S. & Rao, J. M. (2016). 'Filter Bubbles, Echo Chambers, and Online News Consumption'. *Public Opinion Quarterly*, 80 (Special issue): 298–320. http://sethrf.com/files/bubbles.pdf

Fletcher, R., & Nielsen, R. K. (2018). 'Are people incidentally exposed to news on social media? A comparative analysis'. *New Media & Society*, 17 August 2017. 20(7), 2450-2468 https://doi.org/10.1177/1461444817724170

Flore, M., Balahur, A., Podavini, A. & Verile, M. (2019). 'Understanding Citizens' Vulnerabilities to Disinformation and Data-Driven Propaganda'. EUR 29741 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-03320-2, doi:10.2760/919835, JRC116009. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC116009/understanding_citizens_vulnerabilities_to_disinformation.pdf

Foer, F. (2017). 'World Without Mind: The Existential Threat of Big Tech'. Penguin, London.

FOJO: Media Institute. (2018). '#journodefender: Turning trolling against journalists on its head'. October 2018. https://journodefender.org/media/SE_journodefender_public_v1.pdf

Fortune, C. (2018). 'Digitally dissecting atrocities – Amnesty International's open source investigations'. *Amnesty International*, 26 September 2018. https://www.amnesty.org/en/latest/news/2018/09/digitally-dissecting-atrocities-amnesty-internationals-open-source-investigations/

France Fight against Manipulation of Information Law. (2018). 'LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information (1)'. *Legifrance*, 23 December 2018. https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037847559&categorieLien=id

France 24. (2020). 'Misinformation flood hampers fight for virus vaccine in Africa'. 07 May 2020. https://www.france24.com/en/20200507-misinformation-flood-hampers-fight-for-virus-vaccine-in-africa

François, C. (2019). 'Actors, Behaviors, Content: A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry & Regulatory Responses', *Transatlantic Working Group*, 20 September 2019. https://www.ivir.nl/publicaties/download/ABC_Framework_2019_Sept_2019.pdf

François, C., Nimmo, B. & Shawn Eib, C. (2019). 'The IRA CopyPasta Campaign'. *Graphika*, October 2019. https://graphika.com/uploads/Graphika%20Report%20-%20CopyPasta.pdf

Frau-Meigs, D. (2019). 'Information Disorders: Risks and Opportunities for Digital Media and Information Literacy?' Media Studies 10, 19 (2019): 11-27 https://hrcak.srce.hr/ojs/index.php/medijske-studije/issue/view/392

Frau-Meigs, D. (in press). 'Addressing the risks of harms caused by disinformation: European vs American approaches to testing the limits of dignity and freedom of expression online'. In Handbook of Communication Rights, Law & Ethics, L. Corredoira (ed), Wiley, forthcoming.

Freedman, D. (2016). 'Divided Britain? We were already divided...' In EU Referendum Analysis 2016: Media, Voters and the Campaign. Jackson, D., Thorsen, E., Wring, D. Loughborough University Center for the Study of Journalism, Culture and Community. http://www.referendumanalysis.eu/

Free Malaysia Today. (2019). 'Malaysia finally scraps Anti-Fake News Act'. 19 December 2019. https://www.freemalaysiatoday.com/category/nation/2019/12/19/malaysia-finally-scraps-anti-fake-news-act/

French Parliament. (2019). 'Lutte contre la haine sur internet - Proposition de loi. n° 1785 , déposé(e) le mercredi 20 mars 2019'. http://www.assemblee-nationale.fr/dyn/15/dossiers/lutte_contre_haine_internet

Friedman, U. (2020). 'The Coronavirus-Denial Movement Now Has a Leader'. *The Atlantic*, 27 March 2020. https://www.theatlantic.com/politics/archive/2020/03/bolsonaro-coronavirus-denial-brazil-trump/608926/

Fries, F. (2018). 'The Role of a Global News Agency In The Era of Big Tech And Fake News'. *FCC*, 10 December 2018. https://www.fcchk.org/event/club-lunch-the-role-of-a-global-news-agency-in-the-era-of-big-tech-and-fake-news/

Full Fact. (2018). 'Tackling misinformation in an open society'. https://fullfact.org/media/uploads/full_fact_tackling_misinformation_in_an_open_society.pdf

Full Fact (2019). 'Report on the Facebook Third Party Fact Checking programme'. https://fullfact.org/media/uploads/tpfc-q1q2-2019.pdf

Funke. D. (2018). 'Automated fact-checking has come a long way. But it still faces significant challenges'. *Poynter*, 4 April, 2018. https://www.poynter.org/news/automated-fact-checking-has-come-long-way-it-still-faces-significant-challenges

Funke, D. (2019). ''I spent almost a month on a floor': What it's like to be imprisoned on false news charges'. *Poynter*, 22 January 2019. https://www.poynter.org/fact-checking/2019/i-spent-almost-a-month-on-a-floor-what-its-like-to-be-imprisoned-on-false-news-charges/

Funke, D. & Benkelman, S. (2019). 'Factually: Games to teach media literacy'. American Press Institute, 18 July 2019. https://www.americanpressinstitute.org/fact-checking-project/factually-newsletter/factually-games-to-teach-media-literacy/

Funke, D. & Mantzarlis, A. (2018b). 'Here's what to expect from fact-checking in 2019'. *Poynter*, 18 December 2018. https://www.poynter.org/fact-checking/2018/heres-what-to-expect-from-fact-checking-in-2019/

Funke, D. & Mantzarlis, A. (2018a). 'We asked 19 fact-checkers what they think of their partnership with Facebook. Here's what they told us'. *Poynter*, 14 December 2018. https://www.poynter.org/fact-checking/2018/we-asked-19-fact-checkers-what-they-think-of-their-partnership-with-facebook-heres-what-they-told-us/

Gadde, V. (2020). 'Protecting and supporting journalists during COVID-19'. *Twitter*, 24 March 2020. https://blog.twitter.com/en_us/topics/company/2020/giving-back-covid-19.html

Gadde, V. & Derella, M. (2020). 'An update on our continuity strategy during COVID-19'. *Twitter Blog*, 16 March 2020. https://blog.twitter.com/en_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html

Galbraith, J. K. (1971). 'A contemporary guide to economics, peace and laughter. Essays edited by Andrea D. Williams'. Chapter 3: How Keynes Came to America, Quote Page 50, Houghton Mifflin Company, Boston, Massachusetts.

Galtung, J. & Ruge, M.H. (1965). 'The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers'. *Journal of Peace Research*, 2(1), 64-90. https://www.semanticscholar.org/paper/The-Structure-of-Foreign-News-The-Presentation-of-Galtung-Ruge/b3b929df1fd2aa3ea6ddd7b44448fd409e48ea0a

Garside, J. (2020). 'Rappler editor Maria Ressa: 'They could shut us down tomorrow''. *The Guardian*, 26 February 2020. https://www.theguardian.com/world/2020/feb/26/rappler-editor-maria-ressa-journalist-they-could-shut-us-down-tomorrow-philippines-fake-news

Gaw, F. (2020). 'Digital disinformation is as potent as a virus during a pandemic'. *Rappler*, 20 March 2020. https://www.rappler.com/technology/features/255224-digital-disinformation-fake-news-coronavirus

Gentzkow, M. & Shapiro, J. M. (2011). 'Ideological segregation online and offline." *Quarterly Journal of Economics*, 126 (4): 1799–1839. (DOI): 10.3386/w15916 https://www.nber.org/papers/w15916

German BMJV. (2020a). 'Gesetzespaket gegen Rechtsextremismus und Hasskriminalität, 2020'. *German Federal Ministry of Justice and Consumer Protection*. 19 February, 2020. https://www.bmjv.de/SharedDocs/Artikel/DE/2020/021920_Kabinett_Bekaempfung_Rechtsextremismus_Hasskriminalitaet.html

German BMJV. (2020b). 'Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität, 2020'. German Federal Ministry of Justice and Consumer Protection. https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/RefE_NetzDGAendG.pdf?__blob=publicationFile&v=3

German NetzDG. (2017). 'Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG)'. *Deutscher Bundestag*. https://dipbt.bundestag.de/extrakt/ba/WP18/815/81582.html

German NetzDG English translation. (2017). 'Network Enforcement Act (Netzdurchsetzunggesetz, NetzDG)' German Law Archive, 1 October 2017. https://germanlawarchive.iuscomp.org/?p=1245

Gettleman, J., Goel, V. & Abi-Habib, M. (2019). 'India Adopts the Tactic of Authoritarians: Shutting Down the Internet'. *The New York Times*, 17 December 2019. https://www.nytimes.com/2019/12/17/world/asia/india-internet-modi-protests.html

Giglietto, F., Iannelli, L., Rossi, L. & Valeriani, A. (2016). 'Fakes, News and the Election: A New Taxonomy for the Study of Misleading Information within the Hybrid Media System'. Convegno AssoComPol 2016. SSRN, https://ssrn.com/abstract=2878774

GIJN Staff. (2019). 'Full Text: Maria Ressa's Keynote Speech for #GIJC19'. 08 October 2019. https://gijn.org/2019/10/08/full-text-maria-ressas-keynote-speech-for-gijc19/

Gilbert, B. (2019). 'Facebook refuses to fact-check political ads, and it's infuriating employees and lawmakers. Here's why the issue continues to dog the company.' *Business Insider*, 14 December 2019. https://www.businessinsider.com/facebook-political-ads-fact-check-policy-explained-2019-11?r=US&IR=T

Gillespie, T. (2017). 'The platform metaphor, revisited'. *Hig Science Blog*, *Institut fur Internet und Gesellschaft*, 24 August 2017. https://www.hiig.de/en/the-platform-metaphor-revisited/

Glazer, E. (2019). 'Facebook Weighs Steps to Curb Narrowly Targeted Political Ads'. *The Wall Street Journal*, 21 November 2019. https://www.wsj.com/articles/facebook-discussing-potential-changes-to-political-ad-policy-11574352887?redirect=amp

Gleicher, N. (2018a). 'Coordinated Inauthentic Behavior Explained'. *Facebook*, 6 December 2018. https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/

Gleicher, N. (2018b). 'How We Work With Our Partners to Combat Information Operations'. *Facebook Newsroom*, 13 November 2018. https://about.fb.com/news/2018/11/last-weeks-takedowns/#working-with-partners

Gleicher, N. (2019). 'How We Respond to Inauthentic Behavior on Our Platforms: Policy Update'. *Facebook Newsroom*, 21 October 2019. https://about.fb.com/news/2019/10/inauthentic-behavior-policy-update/

Gleicher, N. (2020). 'Removing Coordinated Inauthentic Behavior From Russia, Iran, Vietnam and Myanmar'. *Facebook*, 12 February, 2020. https://about.fb.com/news/2020/02/removing-coordinated-inauthentic-behavior/

Global Disinformation Index. (2019). 'The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?' September 2019. https://disinformationindex.org/wp-content/uploads/2019/09/GDI_Ad-tech_Report_Screen_AW16.pdf

Goel, V., Deep Singh, K. & Yasir, S. (2019). 'India Shut Down Kashmir's Internet Access. Now, 'We Cannot Do Anything.'. *The New York Times*, 14 August 2019. https://www.nytimes.com/2019/08/14/technology/india-kashmir-internet.html

Goggin, B. & Tenbarge, K. (2019). ''Like you've been fired from your job': YouTubers have lost thousands of dollars after their channels were mistakenly demonetized for months'. *Business Insider*, 24 August 2019. https://www.businessinsider.com/youtubers-entire-channels-can-get-mistakenly-demonetized-for-months-2019-8?r=US&IR=T

Goldshlager, K. & Watson, O. (2020). 'Launching a $1M Grant Program to Support Fact-Checkers Amid COVID-19'. *Facebook Journalism Project*, 30 April 2020. https://www.facebook.com/journalismproject/coronavirus-grants-fact-checking

Goldzweig, R. (2020). 'It is time tech companies act on election-time disinformation'. *Al Jazeera*, 23 May 2020. https://www.aljazeera.com/indepth/opinion/time-tech-companies-act-election-time-disinformation-200520135809708.html

Google (2019). Political Advertising on Google. https://transparencyreport.google.com/political-ads/home

Google (2019). YouTube Community Guidelines Enforcement. https://transparencyreport.google.com/youtube-policy/removals

Google and YouTube (2019). 'EC EU Code of Practice on Disinformation – Google Annual Report', https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019

Gorwa, R., Binns, R. & Katzenbach, C. (2020). 'Algorithmic content moderation: Technical and political challenges in the automation of platform governance', *Big Data & Society*, January - June 2020, 1-15. https://doi.org/10.1177/2053951719897945

Gottfried, J. & Greco, E. (2018). 'Younger Americans are better than older Americans at telling factual news statements from opinions'. *Pew Research Center*, 23 October 2018. http://www.pewresearch.org/fact-tank/2018/10/23/younger-americans-are-better-than-older-americans-at-telling-factual-news-statements-from-opinions/,

Gottfried, J., Barthel, M. & Mitchell, A. (2017). 'Trump, Clinton Voters Divided in Their Main Source for Election News'. *Pew Research Centre*, 18 January 2017. https://www.journalism.org/2017/01/18/trump-clinton-voters-divided-in-their-main-source-for-election-news/

Government of Canada. (2019a). 'Response to Canadian Parliamentary Standing Committee on Access to Information, Privacy and Ethics Report. Democracy under Threat: Risks and Solutions in the Era of Disinformation and Data Monopoly'. https://www.ourcommons.ca/DocumentViewer/en/42-1/ETHI/report-17/response-8512-421-502

Government of Canada. (2019b). 'Online Disinformation'. https://www.canada.ca/en/canadian-heritage/services/online-disinformation.html

Graells-Garrido, E., Lalmas, M. & Baeza-Yates, R. (2016). 'Data Portraits and Intermediary Topics: Encouraging Exploration of Politically Diverse Profiles'. *In Proceedings of the 21st International Conference on Intelligent User Interfaces*, March 2016 https://doi.org/10.1145/2856767.2856776

Grau, M. (2020). 'New WhatsApp chatbot unleashes power of worldwide fact-checking organizations to fight COVID-19 misinformation on the platform'. *Poynter*, 04 May 2020. https://www.poynter.org/fact-checking/2020/poynters-international-fact-checking-network-launches-whatsapp-chatbot-to-fight-covid-19-misinformation-leveraging-database-of-more-than-4000-hoaxes/

Graves, L. (2013). 'Deciding What's True: Fact-Checking Journalism and the New Ecology of News'. Columbia University. Philosophy PhD dissertation. https://core.ac.uk/download/pdf/161442732.pdf

Graves, L. (2018). 'FACTSHEET: Understanding the Promise and Limits of Automated Fact-Checking'. *Reuters*. http://www.digitalnewsreport.org/publications/2018/factsheet-understanding-promise-limits-automated-fact-checking/

Graves, L, & Cherubini, F. (2016). 'The Rise of Fact-Checking Sites in Europe'. *Reuters*, http://www.digitalnewsreport.org/publications/2016/rise-fact-checking-sites-europe/

Green, V. (2019). 'Our staff was clear: Facebook's money isn't worth it'. https://twitter.com/vinnysgreen/status/1187135050761920512

Green, V. & Mikkelson, D. (2019). 'A Message to Our Community Regarding the Facebook Fact-Checking Partnership'. *Snopes*, 1 February 2019. https://www.snopes.com/blog/2019/02/01/snopes-fb-partnership-ends/

Gregory, S. (2019). 'Deepfakes and Synthetic Media: Updated Survey of Solutions against Malicious Usages'. *Witness*, June 2019. https://blog.witness.org/2019/06/deepfakes-synthetic-media-updated-survey-solutions-malicious-usages/

Gregory, S. & French, E. (2019). 'How do we work together to detect AI-manipulated media?'. *Witness Media Lab*. https://lab.witness.org/projects/osint-digital-forensics/

Grigoryan, A. (2019). 'Initiatives to Counter Fake News: Russia'. *Library on Congress*, April 2019. https://www.loc.gov/law/help/fake-news/russia.php#_ftn17

Grose, C. R. & Husser, J.A. (2008). 'The Valence Advantage of Presidential Persuasion: Do Presidential Candidates Use Oratory to Persuade Citizens to Vote Contrary to Ideological Preferences?' 03 April 2008. https://ssrn.com/abstract=1119444 or http://dx.doi.org/10.2139/ssrn.1119444

Grossman, D. & Schickler, R. (2019). 'Facebook took down our fact-check on medically necessary abortions. That's dangerous.' *The Washington Post*, 15 September 2019. https://www.washingtonpost.com/opinions/2019/09/15/facebook-took-down-our-fact-check-medically-necessary-abortions-thats-dangerous/

Guadagno, R. E., & Guttieri, K. (2019). 'Fake News and Information Warfare: An Examination of the Political and Psychological Processes From the Digital Sphere to the Real World'. In *Handbook of Research on Deception, Fake News, and Misinformation Online* (pp. 167-191). IGI Global

Guess, A., Nagler, J. and Tucker, J. (2019). 'Less than you think: Prevalence and predictors of fake news dissemination on Facebook'. *Science Advances*, Vol. 5 (1), p.eaau4586. https://advances.sciencemag.org/content/5/1/eaau4586

Guess, A., Nyhan, B., Lyons, B., & Reifler, J. (2018a). 'Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think'. Knight Foundation White Paper. https://www.researchgate.net/publication/330144926_Avoiding_the_echo_chamber_about_echo_chambers_Why_selective_exposure_to_like-minded_political_news_is_less_prevalent_than_you_think

Guess, A., Nyhan, B., Reifler, J. (2018b). "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 US. Presidential Campaign". http://www.dartmouth.edu/~nyhan/fake-news-2016.pdf

Gupta, S. (2020). 'Journalism in the time of corona: This is the biggest story of our lives'. *The Print*, 21 March 2020. https://theprint.in/national-interest/journalism-in-the-time-of-corona-this-is-the-biggest-story-of-our-lives/385057/?amp&__twitter_impression=true

Haffajee, F. & Davies, M. (2017). 'Ferial Haffajee: The Gupta Fake News Factory And Me'. *Huffington Post*, 06 June 2017. https://www.huffingtonpost.co.uk/2017/06/05/ferial-haffajee-the-gupta-fake-news-factory-and-me_a_22126282/

Haffajee, F. (2019). '#GuptaLeaks wins Global Shining Light investigative journalism award'. *Daily Maverick*, 29 September 2019. https://www.dailymaverick.co.za/article/2019-09-29-guptaleaks-wins-global-shining-light-investigative-journalism-award/

Halon, Y. (2020). 'Zuckerberg knocks Twitter for fact-checking Trump, says private companies shouldn't be 'the arbiter of truth'. *Fox News*, 28 May 2020. https://www.foxnews.com/media/facebook-mark-zuckerberg-twitter-fact-checking-trump

Hamborg, F., Donnay, K. & Gipp, B. (2018). 'Automated identification of media bias in news articles: an interdisciplinary literature review'. 16 November 2018. Int J Digit Libr 20, 391–415 (2019). https://doi.org/10.1007/s00799-018-0261-y

Hanly, K. (2018). 'Op-Ed: Linguist George Lakoff explains how Trump tactics work'. Digital Journal, 17 November 2018. http://www.digitaljournal.com/news/politics/op-ed-linguist-george-lakoff-explains-how-trump-tactics-work/article/537117

Hansen, F. S. (2017). 'The weaponization of information'. *DIIS*, 14 December 2017. https://www.diis.dk/en/research/the-weaponization-of-information

Hao, K. (2019). 'This is how AI bias really happens—and why it's so hard to fix'. *MIT Technology Review*, 04 February 2019. https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/

Hanusch, F. (2017). 'Web analytics and the functional differentiation of journalism cultures: Individual, organizational and platform-specific influences on newswork'. Information, Communication & Society, 20(10), 1571-1586. https://www.tandfonline.com/doi/full/10.1080/1369118X.2016.1241294

Harding McGill, M. & Daly, K. (2020). 'New calls to curtail tech's targeted political advertising'. *Axios*, 28 May 2020. https://www.axios.com/new-calls-to-curtail-techs-targeted-political-advertising-9dcd7d8a-4d27-4846-92a1-6f4aab42d7b2.html

Harris, B. (2019). 'An Update on Building a Global Oversight Board'. *Facebook Newsroom*, 12 December 2019. https://about.fb.com/news/2019/12/oversight-board-update/

Harvey, D. (2019). 'Helping you find reliable public health information on Twitter'. *Twitter Blog*, 10 May 2019. https://blog.twitter.com/en_us/topics/company/2019/helping-you-find-reliable-public-health-information-on-twitter.html

Harvey, D. & Roth, Y. (2018). 'An update on our elections integrity work'. *Twitter*, 1 October 2018. https://blog.twitter.com/official/en_us/topics/company/2018/an-update-on-our-elections-integrity-work.html

Hatmaker, T. (2020). 'Jack Dorsey explains why Twitter fact-checked Trump's false voting claims'. *Techcrunch*, 28 May 2020. https://techcrunch.com/2020/05/27/twitter-vs-trump-fact-checking-dorsey/

Hazard Owen, L. (2019). 'Full Fact has been fact-checking Facebook posts for six months. Here's what they think needs to change'. *Nieman Lab*, 29 July 2019. https://www.niemanlab.org/2019/07/full-fact-has-been-fact-checking-facebook-posts-for-six-months-heres-what-they-think-needs-to-change/

Henderson, A. (2020). 'NYT slammed for 'terminal both-sides-ism' after reporting on Trump's household disinfectant suggestion'. *AlterNet*, 24 April 2020. https://www.alternet.org/2020/04/nyt-slammed-for-terminal-both-sides-ism-after-reporting-on-trumps-household-disinfectant-suggestion/

Henley, J. (2020). 'How Finland starts its fight against fake news in primary schools'. *The Guardian*, 29 January 2020. https://www.theguardian.com/world/2020/jan/28/fact-from-fiction-finlands-new-lessons-in-combating-fake-news

Hern, A. (2019a). 'Facebook's only Dutch factchecker quits over political ad exemption'. *The Guardian*, 27 November 2019. https://www.theguardian.com/technology/2019/nov/27/facebook-only-dutch-factchecker-quits-over-political-ad-exemption

Hern, A. (2019b) 'Facebook fact checkers did not know they could vet adverts'. *The Guardian*, 26 October 2019. https://www.theguardian.com/technology/2019/oct/26/facebook-fact-checkers-paid-adverts-misinformation-mark-zuckerberg-congress

Hern, A. (2020). 'YouTube ads of 100 top brands fund climate misinformation – study'. *The Guardian*, 16 January 2020. https://www.theguardian.com/technology/2020/jan/16/youtube-ads-of-100-top-brands-fund-climate-misinformation-study

Hogan, L. (2018). 'Myanmar groups criticise Zuckerberg's response to hate speech on Facebook'. *The Guardian*, 05 April 2018. https://www.theguardian.com/technology/2018/apr/06/myanmar-facebook-criticise-mark-zuckerberg-response-hate-speech-spread

Hoggins, T. (2019). 'Google and Facebook's moderators are superheroes, but they need more protection'. *The Telegraph*, 17 December 2019. https://www.telegraph.co.uk/technology/2019/12/17/google-facebooks-moderators-superheroes-need-protection/

Hollowood, E. & Mostrous, A. (2020). 'Fake news in the time of C-19: From miraculous cures to paranoid conspiracies, our investigation reveals how misinformation about coronavirus is going viral at a disturbing rate'. *Tortoise*, 23 March 2020. https://members.tortoisemedia.com/2020/03/23/the-infodemic-fake-news-coronavirus/content.html

Hollyfield, A. (2013). 'PolitiFact to launch PunditFact, checking pundits and media figures'. *Politifact*, 10 October 2013. https://www.politifact.com/article/2013/oct/10/politifact-launch-punditfact-checking-pundits-and-/

Horwitz, J. & Seetharaman, D. (2020). 'Facebook Executives Shut Down Efforts to Make the Site Less Divisive'. *The Wall Street Journal*, 26 May 2020. https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499?utm_source=Daily+Lab+email+list&utm_campaign=9e6150a2bf-dailylabemail3&utm_medium=email&utm_term=0_d68264fd5e-9e6150a2bf-396529883

Houngbadji, C. S. (2020). 'Bénin: après Ignace Sossou, un autre journaliste fait les frais du code du numérique'. *Benin Web TV*, 07 January 2020. https://beninwebtv.com/2020/01/benin-apres-ignace-sossou-un-autre-journaliste-fait-les-frais-du-code-du-numerique/

Howard, P. & Bell, E. (2020). 'Disinformation in 2020, from "Plandemic" to Bill Gates to "Obamagate". Interview with Recode Decode's Kara Swisher, 27 May 2020. https://www.podchaser.com/podcasts/recode-decode-100800/episodes/phil-howard-and-emily-bell-dis-61618752

Howard, P. N., Ganesh, B., Liotsiou, D., Kelly, J. & François, C. (2018). 'The IRA, Social Media and Political Polarization in the United States, 2012-2018'. *Computational Propaganda Research Project*. https://comprop.oii.ox.ac.uk/research/ira-political-polarization/

Human Rights Watch. (2018a). 'Kazakhstan: Criminal Probe of Media Outlets'. 06 April 2018. https://www.hrw.org/news/2018/04/06/kazakhstan-criminal-probe-media-outlets#

Human Rights Watch. (2018b). 'Germany: Flawed Social Media Law'. 14 February 2018. https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law

Humprecht, E., Esser, F. & Van Aelst P. (2020). 'Resilience to Online Disinformation: A Framework for Cross-National Comparative Research'. 24 January 2020. https://doi.org/10.1177/1940161219900126

IAMAI. (2019). 'Voluntary Code of Ethics for the General Elections 2019'. http://164.100.117.97/WriteReadData/userfiles/Voluntary%20Code%20of%20Ethics%20for%20the%20G.E.%202019.pdf.

IFCN. (2019a). The International Fact-Checking Network, *Poynter*. https://www.poynter.org/ifcn/

IFCN. (2019c). 'The commitments of the code of principles'. https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles. Visited in December 2019.

IFCN. (2019d). 'About the International Fact-Checking Network'. https://www.poynter.org/about-the-international-fact-checking-network/. Visited in December 2019.

IFCN. (2020a). 'IFCN receives $1 million from WhatsApp to support fact-checkers on the coronavirus battlefront'. *Poynter*, 18 March 2020. https://www.poynter.org/fact-checking/2020/ifcn-receives-1-million-from-whatsapp-to-support-fact-checkers-on-the-coronavirus-battlefront/

IFCN. (2020b). 'Flash grants of up to $50K are now available for fact-checkers fighting coronavirus misinformation'. *Poynter*, 17 March 2020. https://www.poynter.org/fact-checking/2020/flash-grants-of-up-to-50k-are-now-available-for-fact-checkers-fighting-coronavirus-misinformation/

IFCN. (2020c). 'The Coronavirus Fact-Checking Grants will support 13 projects on its 1st round: More than half a million dollars is being distributed'. *Poynter*, 02 April 2020. https://www.poynter.org/fact-checking/2020/the-coronavirus-fact-checking-grants-will-support-13-projects-on-its-1st-round-more-than-half-a-million-dollars-is-being-distributed/

IFCN. (2020d). 'Verified signatories of the IFCN code of principles'. https://ifcncodeofprinciples.poynter.org/signatories. Visited on 6 August 2020.

Indian (MeitY) Government (2018). 'The Information Technology [Intermediaries Guidelines (Amendment) Rules] 2018' https://www.meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf

Indian Ministry of Communications. (2017). 'Temporary Suspension of Telecom Services Rules'. https://dot.gov.in/sites/default/files/Suspension%20Rules.pdf?download=1

Ingber, S. (2019). 'Students in Ukraine Learn How To Spot Fake Stories, Propaganda And Hate Speech'. *NPR*, 22 March 2019. https://www.npr.org/2019/03/22/705809811/students-in-ukraine-learn-how-to-spot-fake-stories-propaganda-and-hate-speech?t=1581335026341

Ingram, M. (2018). 'Facebook slammed by UN for its role in Myanmar genocide'. *CJR*, 08 November 2018. https://www.cjr.org/the_media_today/facebook-un-myanmar-genocide.php

Ingram, M. (2019). 'YouTube is all over the map when it comes to offensive content'. *CJR*, 6 June 2019. https://www.cjr.org/the_media_today/youtube-maza-nazis.php

Ireland House of the Oireachtas. (2019). *International Grand Committee on Disinformation and 'Fake News' Dublin, Ireland – Wednesday 6th and Thursday 7th November 2019*. Press Release, 25 October 2019. https://www.oireachtas.ie/en/press-centre/press-releases/20191025-international-grand-committee-on-disinformation-and-fake-news-dublin-ireland-wednesday-6th-and-thursday-7th-november-2019/

Ireland IDG, (2019). 'Government of Ireland, Interdepartmental Group on Security of Ireland's Electoral Process and Disinformation – Progress Report'. https://assets.gov.ie/39188/8c7b6bc1d0d046be915963abfe427e90.pdf

Ireton C. & Posetti, J., eds. (2018). 'Journalism, Fake News' & Disinformation: A Handbook for Journalism Education and Training', pp 21-22. *UNESCO*. https://en.unesco.org/fightfakenews

Irish Department of the Taoiseach. (2019). 'Proposal to Regulate Transparency of Online Political Advertising'. Press Release, 05 November 2019. https://www.gov.ie/en/news/9b96ef-proposal-to-regulate-transparency-of-online-political-advertising/

Irish IDG. (2019). 'Government of Ireland - Interdepartmental Group on Security of Ireland's Electoral Process and Disinformation – Progress Report'. November, 2019. https://assets.gov.ie/39188/8c7b6bc1d0d046be915963abfe427e90.pdf

ISD. (2019). '2019 EU Elections Information Operations Analysis'. https://www.isdglobal.org/wp-content/uploads/2019/05/ISD-EU-Elections-Computational-Propaganda-Analysis_May-24.docx.pdf

ISD. (2020a). 'Covid-19 Disinformation Briefing No.1'. 27 March 2020. https://www.isdglobal.org/wp-content/uploads/2020/03/COVID-19-Briefing-Institute-for-Strategic-Dialogue-27th-March-2020.pdf

ISD. (2020b). 'Covid-19 Disinformation Briefing No.2'. 09 April 2020. https://www.isdglobal.org/wp-content/uploads/2020/04/Covid-19-Briefing-PDF.pdf

Italian Ministry of Education. (2017). 'Scuola, Boldrini e Fedeli presentano decalogo anti-bufale Il progetto riguarderà 4,2 milioni di ragazzi'. Press Release, 31 October 2017. https://www.miur.gov.it/web/guest/-/scuola-boldrini-e-fedeli-presentano-decalogo-anti-bufale-il-progetto-riguardera-4-2-milioni-di-ragazzi

Italian AGCOM. (2018). 'Linee guida per la parità di accesso alle piattaforme online durante la campagna elettorale per le elezioni politiche 2018'. https://www.agcom.it/documents/10179/9478149/Documento+generico+01-02-2018/45429524-3f31-4195-bf46-4f2863af0ff6?version=1.0

Itimu, K. (2019). 'Judgement on the Computer Misuse and Cybercrimes Law to be Passed Next Year'. *Techweez*, 24 October 2019. https://techweez.com/2019/10/24/judgement-on-the-computer-misuse-and-cybercrimes-law-to-be-passed-next-year/

Jack, C. (2017). 'Lexicon of Lies: Terms for Problematic Information'. *Data & Society Research*, 09 August 2017. https://datasociety.net/output/lexicon-of-lies/

Jahangir, R. (2020). 'Desi totkas and fake news — a guide to surviving the Covid-19 'infodemic'. DAWN, 28 March 2020. https://www.dawn.com/news/1544256/desi-totkas-and-fake-news-a-guide-to-surviving-the-covid-19-infodemic

Jamison, A., Broniatowski, D., Dredze, M., Wood-Doughty, Z., Khan, D. & Crouse Quinn, S. (2019). 'Vaccine-related advertising in the Facebook Ad Archive'. *Science Direct*, Vol 38 :3, 16 January 2020. https://www.sciencedirect.com/science/article/pii/S0264410X1931446X?via%3Dihub

Jankowicz, N. (2019). 'Ukraine's Election Is an All-Out Disinformation Battle'. *The Atlantic*, 17 April 2019. https://www.theatlantic.com/international/archive/2019/04/russia-disinformation-ukraine-election/587179/

Japanese Ministry of Internal Affairs and Communications. (2018). 「プラットフォームサービスに関する研究会における検討アジェンダ(案)」に対する提案募集. Press Release, 19 October 2018. https://www.soumu.go.jp/menu_news/s-news/01kiban18_01000052.html

Jeangène Vilmer, J.-B., Escorcia, A., Guillaume, M., & Herrera, J. (2018). 'Information Manipulation: A Challenge for Our Democracies'. A report by the Policy Planning Staff (CAPS Ministry for Europe and Foreign Affairs) and the Institute for Strategic Research (IRSEM, Ministry for the Armed Forces). *Resource Centre on Media Freedom in Europe*, August 2018. https://www.rcmediafreedom.eu/Publications/Reports/Information-Manipulation-A-Challenge-for-Our-Democracies

Jiahao, W. (2019). '防制不實訊息 臉書LINE等5大業者帶頭自律'. *CNA*, 21 June 2019. https://www.cna.com.tw/news/firstnews/201906210183.aspx

Jiji. (2019). 'Japanese panel wants to establish team to fight fake news, with help from U.S. tech giants'. *The Japan Times*, 30 November 2019. https://www.japantimes.co.jp/news/2019/11/30/business/japan-fake-news-gafa/#.XkEz40HgouV

Jin, K-X. (2020). 'Keeping People Safe and Informed About the Coronavirus'. *Facebook Newsroom*, 04 May 2020. https://about.fb.com/news/2020/05/coronavirus/#joint-statement

Johnson, H. M. & Seifert, C. M. (1994). 'Sources of the continued influence effect: When misinformation in memory affects later inferences'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1420–1436. https://doi.org/10.1037/0278-7393.20.6.1420

Kafka, P. 'Rupert Murdoch wanted Mark Zuckerberg to pay him for news stories — and now Facebook is going to do just that'. *Vox*, 24 October 2019. https://www.vox.com/recode/2019/10/24/20929919/facebook-zuckerberg-murdoch-news-publishers-pay-content

Kahn, B. (2019). 'This Fake Green New Deal Ad Perfectly Illustrates Facebook's Bullshit Political Ad Policy [Updated]. *Gizmodo*, 25 October 2019. https://earther.gizmodo.com/this-fake-green-new-deal-ad-perfectly-illustrates-faceb-1839364467

Kajimoto, M. (2018). 'In East and Southeast Asia, misinformation is a visible and growing concern'. *Poynter*, 14 March 2018. https://www.poynter.org/fact-checking/2018/in-east-and-southeast-asia-misinformation-is-a-visible-and-growing-concern/

Kajimoto, M., Stanley S. Editors. (2019). 'Information Disorder in Asia and the Pacific'. *The University of Hong-Kong, Journalism and Media Studies Center*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3134581

Kalogeropoulos, A., Cherubini, F., & Newman, N. (2016). 'The future of online news video'. *Digital News Project*. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/The%2520Future%2520of%2520Online%2520News%2520Video.pdf

Kao, J. (2020). 'How China Built a Twitter Propaganda Machine Then Let It Loose on Coronavirus'. *ProPublica*, 26 March 2020. https://www.propublica.org/article/how-china-built-a-twitter-propaganda-machine-then-let-it-loose-on-coronavirus

Karanbir Gurung, S. (2019). 'Defence ministry approves information warfare branch for Indian army'. *The Economic Times*, 09 march 2019. https://economictimes.indiatimes.com/news/defence/defence-ministry-approves-information-warfare-branch-for-indian-army/articleshow/68329797.cms

Karimi, N. & Gambrell, J. (2020). 'In Iran, false belief a poison fights virus kills hundreds'. *AP News*, 27 March 2020. https://apnews.com/6e04783f95139b5f87a5febe28d72015

Kaur, K., Nair, S., Kwok, Y., Kajimoto, M, Chua, Y. T., Labiste, M. D., Soon, C., Jo, H., Lin, L., Le, T. T. & Kruger, A. (2018). Chapter on India in 'Information Disorder in Asia and the Pacific: Overview of Misinformation Ecosystem in Australia, India, Indonesia, Japan, the Philippines, Singapore, South Korea, Taiwan, and Vietnam'. 10 October 2018. http://dx.doi.org/10.2139/ssrn.3134581

Kaye, D. (2018). 'Report of the special rapporteur on promotion and protection of the right to freedom of opinion and expression'. *Freedex*. https://freedex.org/a-human-rights-approach-to-platform-content-regulation/

Kaye, D. (2020a). 'Disease pandemics and the freedom of opinion and expression'. 23 April 2020. https://freedex.org/wp-content/blogs.dir/2015/files/2020/04/A_HRC_44_49_AdvanceEditedVersion.pdf

Kaye, D. (2020b). Transcript of interview with Recode Decode's Kara Swisher. https://docs.google.com/document/d/1zYuYMTmvoHu_dgajRTSORZpkY1DhiN8Kb9KytCgVU7g/edit Recode Decode episode: S1, E480, 14 February 2020.

Kazakhstan Penal Code. (2014). https://www.unodc.org/res/cld/document/penal-code_html/New_penal_code.pdf

Keane, J. (2018). 'Post-truth politics and why the antidote isn't simply 'fact-checking' and truth'. *The Conversation*, 23 March 2018. https://theconversation.com/post-truth-politics-and-why-the-antidote-isnt-simply-fact-checking-and-truth-87364

Keller, I. (2019). 'A school of bitter experience: how Kazakhstan's media regulations restrict journalist freedom'. *Open Democracy*, 6 June 2019. https://www.opendemocracy.net/en/odr/press-freedom-kazakhstan-en/

Keller, F. B., Schoch, D., Stier, S & Yang, J. H. (2019). "Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign." Political Communication (2019): 1-25. https://www.tandfonline.com/doi/full/10.1080/10584609.2019.1661888

Kelly, J. & François, C. (2018). 'This is what filter bubbles actually look like'. *MIT Technology Review*. https://www.technologyreview.com/s/611807/this-is-what-filter-bubbles-actually-look-like/

Kelly, M. (2020a). 'The World Health Organization has joined TikTok to fight coronavirus misinformation'. *The Verge*, 28 February 2020. https://www.theverge.com/2020/2/28/21158276/coronavirus-covid19-tiktok-who-world-health-organization-protection

Kelly, M. (2020b). 'Democrats want to restrict political ad targeting ahead of the 2020 election'. *The Verge*, 26 May 2020. https://www.theverge.com/2020/5/26/21271074/facebook-google-microtargeting-political-ads-ban-anna-eshoo

Kelly Garrett, R. (2009). 'Echo chambers online?: Politically motivated selective exposure among Internet news users'. *Journal of Computer-Mediated Communication*, 14(2), 265-285, 30 March 2019. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2009.01440.x

Kenya Computer Misuse and Cybercrimes Act. (2018). http://kenyalaw.org/kl/fileadmin/pdfdownloads/Acts/ComputerMisuseandCybercrimesActNo5of2018.pdf

Kessler, G. (2017). 'Washington Post: About the Fact Checker'. https://www.washingtonpost.com/politics/2019/01/07/about-fact-checker/

Kiesel, J., Mestre, M., Shukla, R., Vincent, E., Adineh, P., Corney, D., Stein, B. & Potthast, M. (2019). 'SemEval-2019 Task 4: Hyperpartisan News Detection'. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, June 2019. (ππ. 829-839). ΔOI: ̈10.18653/ω1/Σ19-2145 https://www.aclweb.org/anthology/S19-2145

King, G. & Persily, N. (2020). 'Unprecedented Facebook URLs Dataset now Available for Academic Research through Social Science One'. *Social Science One*, 13 February 2020. https://socialscience.one/blog/unprecedented-facebook-urls-dataset-now-available-research-through-social-science-one

Khidhir, S. (2019). 'Indonesia has a fake news problem'. *The Asean Post*, 19 November 2019. https://theaseanpost.com/article/indonesia-has-fake-news-problem

Kleinman, Z. (2016). 'Fury over Facebook 'Napalm girl' censorship'. *BBC*, 09 September 2016. https://www.bbc.co.uk/news/technology-37318031#:~:text=Facebook%20said%20it%20has%20to,and%20before%20it%20had%20responded

Knight Foundation. (2018). 'In the internet we trust: the impact of engaging with news articles'. https://knightfoundation.org/reports/in-the-internet-we-trust-the-impact-of-engaging-with-news-articles/

Knockel, J. & Xiong, R. (2019). '(Can't) Picture This 2: An Analysis of WeChat's Realtime Image Filtering in Chats'. *The Citizen Lab*, 15 July 2019. https://citizenlab.ca/2019/07/cant-picture-this-2-an-analysis-of-wechats-realtime-image-filtering-in-chats/

Kongkea, B. R. (2019a). 'Man charged over YouTube fake news'. *Khmer Times*, 02 April 2019. https://www.khmertimeskh.com/592313/man-charged-over-airing-of-fake-news-on-youtube/

Kongkea, B. R. (2019b). 'Former monk convicted over fake news'. *Khmer Times*, 21 October 2019. https://www.khmertimeskh.com/50652677/former-monk-convicted-over-fake-news

Kovach, B. & Rosenstiel, T. (2001). 'The Elements of Journalism, Crown publishers'. New York.

Kozlowska, H. (2019). 'TikTok banned political ads—but pro-Trump content is thriving and misleading teens'. *Quartz*, 24 October 2019. https://qz.com/1731170/pro-trump-videos-are-thriving-on-tiktok/

Kroes, R. (2012). 'The power of rhetoric and the rhetoric of power: Exploring a tension within the Obama presidency'. *European journal of American studies*, *7*(7-2). https://journals.openedition.org/ejas/9578

Kuczerawy, A. (2019). 'Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression?'. Forthcoming in: "Disinformation and Digital Media as a Challenge for Democracy" European Integration and Democracy Series, 6. https://ssrn.com/abstract=3453732

Kutty, S. (2018). 'Fake News - Jail and fine for spreading false news'. *Oman Daily Observer*, 25 May 2018. https://www.omanobserver.om/jail-and-fine-for-spreading-false-news/

La Cour, C. (2019). 'Governments Countering Disinformation: The Case of Italy'. *Disinfo Portal*, 20 November 2019. https://disinfoportal.org/governments-countering-disinformation-the-case-of-italy/

Lacy, L. & Rosenstiel, T. (2015). 'Defining and measuring quality journalism'. *Rutgers*, March 2015. https://www.issuelab.org/resources/31212/31212.pdf

Lamb, K. (2018). 'Cambodia 'fake news' crackdown prompts fears over press freedom'. *The Guardian*, 6 July 2018. https://www.theguardian.com/world/2018/jul/06/cambodia-fake-news-crackdown-prompts-fears-over-press-freedom?CMP=share_btn_tw

Lapowsky, I. (2018). 'Inside the Research Lab Teaching Facebook About Its Trolls'. *Wired*, 15 August 2018. https://www.wired.com/story/facebook-enlists-dfrlab-track-trolls/

Larson, H. J. (2018). 'The biggest pandemic risk? Viral misinformation'. *Natureresearch*, 16 October 2018. Nature 562, 309 (2018). doi: 10.1038/d41586-018-07034-4 https://www.nature.com/articles/d41586-018-07034-4

Lasica, J. D. (2003). 'Blogs and journalism need each other'. *Nieman reports*, 57(3), 70-74. 15 September 2003. https://niemanreports.org/articles/blogs-and-journalism-need-each-other/

Latvian Public Broadcasting. (2016). 'Latvia shuts down Sputnik propaganda website'. LSM.LV, 29 March 2016. https://eng.lsm.lv/article/society/society/latvia-shuts-down-sputnik-propaganda-website.a175627/

Lazer, D., Baum, M.,Grinberg, N., Friedland, L., Joseph, K., Hobbs, W., & Mattsson, C. (2017). 'Combating Fake News: An Agenda for Research and Action'. https://shorensteincenter.org/combating-fake-news-agenda-for-research/

Leathern, R. (2020). 'Expanded Transparency and More Controls for Political Ads'. *Facebook Newsroom*, 09 January 2020. https://about.fb.com/news/2020/01/political-ads/

Lee, D. (2019a). 'Key fact-checkers stop working with Facebook'. *BBC*, 02 February 2019. https://www.bbc.co.uk/news/technology-47098021

Lee, D. (2019b). 'Matter of fact-checkers: Is Facebook winning the fake news war?' *BBC*, 02 April 2019. https://www.bbc.com/news/technology-47779782

Leerssen, P., Ausloos, J., Zarouali, B., Helberger, N. de Vreese, C. H. (2019). 'Platform ad archives: promises and pitfalls'. *Internet Policy Review*, 8(4). DOI: 10.14763/2019.4.1421 https://policyreview.info/articles/analysis/platform-ad-archives-promises-and-pitfalls

LeFigaro. (2019). 'Loi contre les «fake news»: Twitter bloque une campagne du gouvernement'. *Économie*, 02 April 2019. https://www.lefigaro.fr/flash-eco/loi-contre-les-fake-news-twitter-bloque-une-campagne-du-gouvernement-20190402

Legum, J. (2019). 'Facebook says Trump can lie in his Facebook ads'. *Popular Information*, 3 October 2019. https://popular.info/p/facebook-says-trump-can-lie-in-his

Leibovich, A. (2019). 'Dark Marketing 101: The Change to How Brands Communicate Online'. *Target Marketing*, 26 February 2019. https://www.targetmarketingmag.com/article/dark-marketing-101-change-how-brands-communicate-online/

Leibowicz, C. (2019). 'Protecting Public Discourse from AI-Generated Mis/Disinformation'. Partnership on AI blog, 17 June 2019. https://www.partnershiponai.org/protecting-public-discourse-from-ai-generated-misdisinformation/

Leveson, Lord Justice. (2012). 'An Inquiry into the culture, practices and ethics of the Press'. HC, November 2012. https://www.gov.uk/government/publications/leveson-inquiry-report-into-the-culture-practices-and-ethics-of-the-press

Levien, R. & Aiken, A. (1998). 'Attack resistant trust metrics for public key certification'. In the 7th USENIX Security Symposium, San Antonio, Texas, January 1998. https://www.usenix.org/legacy/publications/library/proceedings/sec98/full_papers/levien/levien.pdf

Levin, A. (2018). ''They don't care': Facebook factchecking in disarray as journalists push to cut ties'. *The Guardian*, 13 December 2018. https://www.theguardian.com/technology/2018/dec/13/they-dont-care-facebook-fact-checking-in-disarray-as-journalists-push-to-cut-ties

Levush, R. (2019). 'Initiatives to Counter Fake News: Israel'. *Library of Congress*, April 2019. https://www.loc.gov/law/help/fake-news/israel.php

Levy, S. (2020). 'Why Mark Zuckerberg's Oversight Board May Kill His Political Ad Policy'. *Wired*, 28 January 2020. https://www.wired.com/story/facebook-oversight-board-bylaws/

Lewis, P. (2018). 'Fiction is outperforming reality': how YouTube's algorithm distorts truth'. *The Guardian*, 02 February 2018. https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth

Liao, Q. V., & Fu, W. T. (2013). 'Beyond the filter bubble: Interactive effects of perceived threat and topic involvement on selective exposure to information'. *In Proceedings of the SIGCHI conference on human factors in computing systems* (*CHI'13*), *2359-2368. ACM*, April 2013. https://doi.org/10.1145/2470654.2481326

Likhachev, N. (2018). 'Как работает политическая реклама во «ВКонтакте»: Собянину можно, Собчак и Навальному — нет'. *TJournal*, 05 January 2018. https://tjournal.ru/news/64619-kak-rabotaet-politicheskaya-reklama-vo-vkontakte-sobyaninu-mozhno-sobchak-i-navalnomu-net

LINE. (2020). '違反投稿への対応'. *Transparency Report*, 15 January 2020. https://linecorp.com/ja/security/moderation/2019h1

LINE. (2019a). 'お役立ち情報をお届け！「スマートチャンネル」を活用してみよう'. 01 April 2019. http://official-blog.line.me/ja/archives/78619192.html

LINE. (2019b). '【公式】LINE広告とは｜サービス概要・特長まとめ'. 24 October 2019. https://www.linebiz.com/jp/column/technique/20191024/

Livingstone, S. (2020). 'Coronavirus and #fakenews: what should families do?' *LSE*, 26 March 2020. https://blogs.lse.ac.uk/medialse/2020/03/26/coronavirus-and-fakenews-what-should-families-do/

Livingstone, S., Byrne, J. & Carr, J. (2016). 'One in Three: Internet Governance and Children's Rights'. *Innocenti Discussion Papers* no. 2016-01, UNICEF Office of Research, January 2016. https://www.unicef-irc.org/publications/795-one-in-three-internet-governance-and-childrens-rights.html

Livsier, L. (2019) 'Ministry to revoke licences of media over fake news'. *Khmer Times*, August 13, 2019. https://www.khmertimeskh.com/632885/ministry-to-revoke-licences-of-media-over-fake-news/

Llansó, E., van Hoboken, J., Leerssen, P. & Harambam, J. (2020). 'Artificial Intelligence, Content Moderation, and Freedom of Expression'. (*Working Paper of the Transatlantic Working Group on Content Moderation Online and Freedom of Expression*). 26 February 2020. https://www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf

Loewenstein, G. (1994). 'The psychology of curiosity: A review and reinterpretation'. *Psychological Bulletin 116*, *1*.

Lomas, N. (2020). 'Facebook's latest 'transparency' tool doesn't offer much — so we went digging'. *Tech Crunch*, 25 February 2020. https://techcrunch.com/2020/02/25/facebooks-latest-transparency-tool-doesnt-offer-much-so-we-went-digging/

Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R., & Hertwig, R. (2020). ,How behavioural sciences can promote truth, autonomy and democratic discourse online'. *Nature Human Behaviour*, 15 June 2020. https://doi.org/10.1038/s41562-020-0889-7

LSE. (2018). 'Tackling the Information Crisis: A Policy Framework for Media System Resilience'. http://www.lse.ac.uk/media-and-communications/assets/documents/research/T3-Report-Tackling-the-Information-Crisis-v6.pdf

Lyons, T. (2018a). 'How is Facebook's Fact-Checking Program Working?' https://newsroom.fb.com/news/2018/06/hard-questions-fact-checking/

Lyons, T. (2018b). 'New Research Shows Facebook Making Strides Against False News'. https://about.fb.com/news/2018/10/inside-feed-michigan-lemonde/

Lypshultz, J. H. (2018). 'Free Expression in the Age of the Internet: Social and Legal Boundaries'. *Routledge*.

MacAvaney, S., Yao, H. R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). 'Hate speech detection: Challenges and solutions'. *PloS one*, 20 August 2019. 14(8). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0221152

MacCarthy, M. (2020). 'Transparency Requirements for Digital Social Media Platforms: Recommendations for Policy Makers and Industry'. One in a Series of Working Papers from the Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression. 12 February 2020. https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/06/Transparency_TWG_MacCarthy_Feb_2020.pdf

MacGuill, D. (2018). 'Did Facebook Flag the Declaration of Independence as Hate Speech?' *Snopes*, 06 July 2018. https://www.snopes.com/fact-check/facebook-declaration-of-independence-hate-speech/

Magdy, S. (2019). 'Egypt tightens restrictions on media, social networks'. *AP*, 19 March 2019. https://apnews.com/1540f1133267485db356db1e58db985b

Malaysian Communications and Multimedia Commission (2020). 'Press Release: Four Individuals Detained For Spreading Fake News On The Novel Coronavirus (2019-nCov) Outbreak'. *Malaysian Communications and Multimedia Commission*, 29 January 2020. https://www.mcmc.gov.my/en/media/press-releases/press-release-four-individuals-detained-for-sprea

Mantzarlis, A. (2017). 'Repetition boosts lies — but could help fact-checkers, too'. *Poynter*, 30 May 2017. https://www.poynter.org/news/repetition-boosts-lies-could-help-fact-checkers-too,

Mantzarlis, A. (2020). 'COVID-19: $6.5 million to help fight coronavirus misinformation'. *Google News Lab*, 02 April 2020. https://www.blog.google/outreach-initiatives/google-news-initiative/covid-19-65-million-help-fight-coronavirus-misinformation/

Maréchal, N. & Biddle, E. R. (2020). 'It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge'. *Open Technology Institute*. https://www.newamerica.org/oti/reports/its-not-just-content-its-business-model/

Maréchal, N., MacKinnon, R. & Biddle, E. R. (2020). 'Getting to the Source of Infodemics: It's the Business Model'. *Open Technology Institute*. https://www.newamerica.org/oti/reports/getting-to-the-source-of-infodemics-its-the-business-model/

Marra, F, Gragnaniello, D, Verdoliva, L. & Poggi, G. (2018). 'Do GANs leave artificial fingerprints?' 31 December 2018. https://arxiv.org/pdf/1812.11842.pdf

Marsden, C. & Meyer, T. (2019). 'Regulating disinformation with artificial intelligence: Effects of disinformation initiatives on freedom of expression and media pluralism. European Parliamentary Research Service, March 2019. https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624279/EPRS_STU(2019)624279_EN.pdf

Martens, B., Aguiar Wicht, L., Gomez-Herrera, M. E. & Mueller-Langer, F. (2018). 'The digital transformation of news media and the rise of disinformation and fake news', *EUR - Scientific and Technical Research Reports, JRC Technical Reports, JRC Digital Economy Working Paper 2018-02*, April 2018. https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digital-transformation-news-media-and-rise-disinformation-and-fake-news#

Martin, D. A., & Shapiro, J. N. (2019). 'Trends in Online Foreign Influence Efforts'. Working Paper, Princeton University, 2019. https://scholar.princeton.edu/sites/default/files/jns/files/trends_in_foreign_influence_efforts_2019jul08_0.pdf

Masih, N., Irfan, S. & Slater, J. (2019). 'India's Internet shutdown in Kashmir is the longest ever in a democracy'. *The Washington Post*, 16 December 2019. https://www.washingtonpost.com/world/asia_pacific/indias-internet-shutdown-in-kashmir-is-now-the-longest-ever-in-a-democracy/2019/12/15/bb0693ea-1dfc-11ea-977a-15a6710ed6da_story.html

Masnick, M. (2003). 'Photo of Streisand Home Becomes An Internet Hit'. *Tech Dirt*, 24 June 2003. https://www.techdirt.com/articles/20030624/1231228.shtml

Mathur, S. (2019). 'I & B team to counter fake news'. *The Times of India*, 16 November 2019. https://timesofindia.indiatimes.com/india/ib-team-to-counter-fake-news/articleshow/72079340.cms?from=mdr&utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst

Mayhew, F. (2020). 'Guardian's head of investigations: 'In times of crisis you need more accountability journalism, not less'. *Press Gazette*, 27 May 2020. https://pressgazette.co.uk/guardian-head-of-investigations-paul-lewis-interview-coronavirus/

Mays, H. (2019). 'Facebook bans Tories' 'distorted' party advert with Laura Kuenssberg'. *The Guardian*, 01 December 2019. https://www.theguardian.com/politics/2019/dec/01/facebook-bans-tories-distorted-party-advert-with-laura-kuenssberg

McLaughlin, T. (2018). 'How WhatsApp Fuels Fake News and Violence in India'. *Wired*, 12 December 2018. https://www.wired.com/story/how-whatsapp-fuels-fake-news-and-violence-in-india/

McNamee, R. (2020). 'Social Media Platforms Claim Moderation Will Reduce Harassment, Disinformation and Conspiracies. It Won't'. *Time*, 24 June 2020. https://time.com/5855733/social-media-platforms-claim-moderation-will-reduce-harassment-disinformation-and-conspiracies-it-wont/

Melo, P. F., Messias, J., Resende, G., Garimella, V. R., Almeida, J. M., & Benevenuto, F. (2019). 'WhatsApp Monitor: A Fact-Checking System for WhatsApp'. In *Proceedings of the Int'l AAAI Conference on Weblogs and Social Media* (*ICWSM'19*). June 2019. https://people.mpi-sws.org/~johnme/pdf/melo-icwsm-2019-demo.pdf

Mexico INE (2019). 'Presenta INE modelo de combate a la desinformación en elecciones 2018-2019'. Press Release, 18 July 2019. https://centralelectoral.ine.mx/2019/07/18/presenta-ine-modelo-combate-la-desinformacion-elecciones-2018-2019/

Mezaris, V. Nixon, L. Papadopoulos, S. & Teyssou, D. (2019). 'Video Verification in the Fake News Era'. Springer.

Milano, B. (2019). 'Israeli Supreme Court Justice on combatting propaganda in elections'. *Harvard Law Today*, 29 October 2019. https://today.law.harvard.edu/israeli-supreme-court-justice-on-combatting-propaganda-in-elections/

Mitchell, A., Gottfried, J., Barthel, M. & Sumida, N. (2018). 'Distinguishing Between Factual and Opinion Statements in the News'. *Pew Research Center*, 18 June, 2018. https://www.journalism.org/wp-content/uploads/sites/8/2018/06/PJ_2018.06.18_fact-opinion_FINAL.pdf

Mitchell, A., Gottfried, J., Stocking, G. Walker, M. & Fedeli, S. (2019). 'Many Americans Say Made-Up News Is a Critical Problem That Needs To Be Fixed'. *Pew Research Centre*, Journalism & Media, 05 June 2019. https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/

Molina, K. (2016). 'Indonesian Electronic Information and Transactions Law Amended'. *White & Case*, 15 December 2016. https://www.whitecase.com/publications/alert/indonesian-electronic-information-and-transactions-law-amended

Monaco, N. & Nyst, C. (2018). 'State-Sponsored Trolling. How Governments Are Deploying Disinformation as Part of Broader Digital Harassment Campaigns'. *IFTF*. http://www.iftf.org/statesponsoredtrolling/

Moore, N. (2020). 'Study Finds More COVID-19 Cases Among Viewers Of Fox News Host Who Downplayed The Pandemic'. *Wbez*, 30 April 2020. https://www.wbez.org/stories/study-finds-more-covid-19-cases-among-viewers-of-fox-news-host-who-downplayed-the-pandemic/83dffd03-6013-47c8-b012-0b1cc5377e18

Moreira, D., Bharati, A., Brogan, J., Pinto, A., Parowski, M., Bowyer, K. W., Flynn, P. J., Rocha, A. & Scheirer, W. J. (2018). 'Image Provenance Analysis at Scale'. 19 January 2018. https://arxiv.org/abs/1801.06510

Mosseri, A. (2016). 'Addressing Hoaxes and Fake News'. https://about.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/

MSB. (2020). 'Countering information influence activities. A handbook for communicators'. Swedish Civil Contingencies Agency. https://rib.msb.se/filer/pdf/28698.pdf

Mullen, A. (2016). 'Leave vs Remain: the Digital Battle' In 'EU Referendum Analysis 2016: Media, Voters and the Campaign'. Jackson, D, Thorsen, E, Wring, D. Loughborough University Center for the Study of Journalism, Culture and Community. http://www.meandeurope.co.uk/wp-content/uploads/EU-Referendum-Analysis-2016-Jackson-Thorsen-and-Wring-v2.pdf

Munson, S. A., Lee, S. Y. & Resnick, P. (2013). 'Encouraging Reading of Diverse Political Viewpoints with a Browser Widget'. In International conference on weblogs and social media (ICWSM). Boston. https://webcache.googleusercontent.com/search?q=cache:TcCE3Axx_9MJ:https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/download/6119/6381+&cd=1&hl=en&ct=clnk&gl=uk&client=firefox-b-e

Myanmar Penal Code. (1861). https://www.burmalibrary.org/docs6/MYANMAR_PENAL_CODE-corr.1.pdf

Myanmar Telecommunications Law. (2013). 'The Telecommunications Law (The Pyidaungsu Hluttaw Law No. 31, 2013) The 4th Waxing Day of Thadingyut, 1375 M.E., 8 October 2013. https://www.burmalibrary.org/docs23/2013-10-08-Telecommunications_Law-en.pdf

Nagulendra, S., & Vassileva, J. (2014). 'Understanding and controlling the filter bubble through interactive visualization: a user study'. *In Proceedings of the 25th ACM conference on Hypertext and social media* (*pp. 107-115*), September 2014. https://www.researchgate.net/publication/266660926_Understanding_and_controlling_the_filter_bubble_through_interactive_visualization_A_user_study

National Literacy Trust. (2018). 'Fake news and critical literacy: final report'. 11 June 2018. https://literacytrust.org.uk/research-services/research-reports/fake-news-and-critical-literacy-final-report/

NATO Stratcom COE (2019). '*About Us*', https://www.stratcomcoe.org/about-us

Neate, R. (2017). 'Bell Pottinger faces hearing over claims it stirred racial tension in South Africa'. *The Guardian*, 13 August 2017. https://www.theguardian.com/world/2017/aug/13/bell-pottinger-pr-industry-hearing-secret-south-africa-campaign

Nechushtai, E., & Lewis, S.C. (2018). 'What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations'. *Computers in Human Behavior*, 90, 298-307. https://doi.org/10.1016/j.chb.2018.07.043

NED. (2018). Comparative Responses to the Global Disinformation Challenge. National Endowment for Democracy. October 4-5, 2018

Nelson, R.A. (1996). 'A Chronology and Glossary of Propaganda in the United States, Westport, Connecticut', p232-233. *Greenwood Press*, ISBN 0313292612

Netherlands Policy Report for Online Disinformation. (2018). *Dropbox*, July, 2018. https://www.dropbox.com/s/99iza9kmbwjbels/20180718_rapport_onlinedesinformatieNL.pdf?dl=0

Newman, L. H. (2019a). 'To Fight Deepfakes, Researchers Built a Smarter Camera'. *Wired*, 28 May 2019. https://www.wired.com/story/detect-deepfakes-camera-watermark/

Newman, N. (2019b). 'Executive Summary and Key Findings of the 2019 Report'. *Reuters, University of Oxford, Digital News Report*. http://www.digitalnewsreport.org/survey/2019/overview-key-findings-2019/

Newman, N. & Fletcher, R. (2017). 'Bias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media'. *Reuters*. https://reutersinstitute.politics.ox.ac.uk/our-research/bias-bullshit-and-lies-audience-perspectives-low-trust-media

Newsguard. (2020a). 'Coronavirus Misinformation Tracking Center'. 20 April 2020. https://www.newsguardtech.com/coronavirus-misinformation-tracking-center/

Newsguard. (2020b). 'NewsGuard Partners with DCMS and BT to Help Counter Spread of COVID-19 Fake News as Misinformation Peaks'. 27 March 2020. https://www.newsguardtech.com/press/newsguard-partners-with-dcms-and-bt-to-help-counter-spread-of-covid-19-fake-news-as-misinformation-peaks/

Newton, C. (2019a). 'Why a top content moderation company quit the business instead of fixing its problems'. *The Verge*, 01 November 2019. https://www.theverge.com/interface/2019/11/1/20941952/cognizant-content-moderation-restructuring-facebook-twitter-google

Newton, C. (2019b). 'The Trauma Floor. The secret lives of Facebook moderators in America'. *The Verge*, 25 February 2019. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona

Newton, C. (2020). 'Snap will stop promoting Trump's account after concluding his tweets incited violence'. *The Verge*, 03 June 2020. https://www.theverge.com/2020/6/3/21279280/snapchat-snap-remove-trump-account-discover-promotion-incite-violence-twitter

New York Times. (2019). 'Read the Letter Facebook Employees Sent to Mark Zuckerberg About Political Ads'. 28 October 2019. https://www.nytimes.com/2019/10/28/technology/facebook-mark-zuckerberg-letter.html

New York Times. (2020). 'How China Is Reshaping the Coronavirus Narrative'. *You Tube*, 18 March 2020. https://www.youtube.com/watch?v=LwqhvRcBrK4&feature=youtu.be

New Zealand Parliament Justice Committee. (2019). Inquiry into the 2017 General Election and 2016 Local Elections. https://www.parliament.nz/resource/en-NZ/SCR_93429/5dd1d57eeba54f36bf9f4da96dba12c073ed7ad8

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology*. 2:175. https://doi.org/10.1037/1089-2680.2.2.175

Nigerian Senate. (2019). Nigeria Protection from Internet Falsehood and Manipulation Bill. (2019). https://ndlink.org/wp-content/uploads/2019/11/Protection-from-Internet-Falsehood-and-Manipulation-Bill-2019.pdf

Nimmo, B. (2019). 'UK Trade Leaks and Secondary Infektion: New Findings and Insights from a Known Russian Operation'. *Graphika*, December 2019. https://graphika.com/uploads/Graphika%20Report%20-%20UK%20Trade%20Leaks%20&%20Secondary%20Infektion.pdf

Nimmo, B., Shawn Eib, C. & Tamora, L. (2019a). 'Cross-Platform Spam Network Targeted Hong Kong Protests. "Spamouflage Dragon" used hijacked and fake accounts to amplify video content'. *Graphika*, September 2019. https://swank-range.cloudvent.net/uploads/Graphika%20Report%20-%20Cross-Platform%20Spam%20Network%20Targeted%20Hong%20Kong%20Protests.pdf

Nimmo, B., Shawn Eib, C., Tamora, L., Johnson, K., Smith, I. (from Graphika), Buziashvili, E., Kann, A., Karan, K., Ponce de Leon Rosas, E. & Rizzuto, M. (from DFRLab) (2019b). '#OperationFFS: Fake Face Swarm'. *Graphika & DFRLab* 20 December 2019. https://graphika.com/reports/operationffs-fake-face-swarm/

Norris, P., Cameron, S., & Wynter, T. (2019). Electoral Integrity in America. Securing Democracy. Oxford University Press.

Nuñez, M. (2020). 'Snap Stock Drops 11% After Revenue Falls Short Of Expectations'. *Forbes*, 04 February 2020. https://www.forbes.com/sites/mnunez/2020/02/04/snap-stock-drops-14-after-revenue-falls-short-of-expectations/#5e7c1bf0448c

Nygren, T. & Guath, M. (2019). 'Swedish teenagers' difficulties and abilities to determine digital news credibility'. *Sciendo*, 19 February 2019. https://doi.org/10.2478/nor-2019-0002

Nygren, T. & Guath, M. (2020). 'A Multidisciplinary Look at Knowledge Resistance'. *YouCheck*, 14 January 2020. http://project-youcheck.com/a-multidisciplinary-look-at-knowledge-resistance/

Nyhan, B. (2012). Columbia Journalism Review: Does Fact-Checking Work? False Statements are Wrong Metric. https://archives.cjr.org/united_states_project/does_fact-checking_work_false.php

Nyhan, B. & Reifler, J. (2006). 'When Corrections Fail: The persistence of political misperceptions' Forthcoming, *Political Behavior*. https://www.dartmouth.edu/~nyhan/nyhan-reifler.pdf.

Nyhan, B & Reifler, J. (2010). 'When Corrections Fail: The Persistence of Political Misperceptions'. *Political Behavior*, 30 March 2010. https://link.springer.com/article/10.1007%2Fs11109-010-9112-2

Nyhan, B.& Reifler, J. (2012). 'Misinformation and Fact-checking: Research Findings from Social Science'. http://web.archive.org/web/20150226014434/http://mediapolicy.newamerica.net/sites/newamerica.net/files/policydocs/Misinformation_and_Fact-checking.pdf

OAS. (2017). 'Office Of The Special Rapporteur For Freedom Of Expression Expresses Serious Concern Over The Enactment Of The "Anti-Hate Law" In Venezuela And Its Effects On Freedom Of Expression And Freedom Of The Press'. Press Release, 10 November 2017. http://www.oas.org/en/iachr/expression/showarticle.asp?artID=1082&lID=1

OAS. (2019). IACHR (Inter-American Commission on Human Rights) & RFOE (Special Rapporteur for Freedom of Expression). Guide to guarantee freedom of expression regarding deliberate disinformation in electoral contexts. http://www.oas.org/en/iachr/expression/publications/Guia_Desinformacion_VF%20ENG.pdf

O'Brien, C. & Kelly, F. (2018). 'Google Bans Online Ads on Abortion Referendum'. *The Irish Times*, 9 May 2018. https://www.irishtimes.com/business/media-and-marketing/google-bans-online-ads-on-abortion-referendum-1.3489046

Ofcom. (2018a). 'News consumption in the UK'. 25 July 2018. https://www.ofcom.org.uk/research-and-data/tv-radio-and-on-demand/news-media/news-consumption.

Ofcom. (2018b). 'Addressing harmful online content'. 18 September 2018. https://www.ofcom.org.uk/phones-telecoms-and-internet/information-for-industry/online-policy-research/addressing-harmful-online-content

Oliphant, R. (2017). 'Ukraine bans Russian social networks in sweeping expansion of sanctions'. *The Telegraph*, 16 May 2017. https://www.telegraph.co.uk/news/2017/05/16/ukraine-bans-russian-social-networks-sweeping-expansion-sanctions/

Ong, J. C. & Cabañes, J. V. A. (2018). 'Architects of Networked Disinformation. Behind the Scenes of Troll Accounts and Fake News Production in the Philippines'. *The Newton Tech4Dev Network*, 17 February 2018. https://newtontechfordev.com/wp-content/uploads/2018/02/ARCHITECTS-OF-NETWORKED-DISINFORMATION-FULL-REPORT.pdf

Ong, J. C. & Cabañes, J. V. A. (2019). 'Politics and Profit in the Fake News Factory: Four Work Models of Political Trolling in the Philippines'. *NATO Strategic Communications Centre of Excellence*, November 2019. https://stratcomcoe.org/four-work-models-political-trolling-philippines.

OSCE. (2017). UN/OSCE/OAS/ACHPR Rapporteurs on Freedom of Expression. (2017). *Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda*'. https://www.osce.org/fom/302796?download=true

Osmundsen, M., Bor, A., Vahlstrup, P.B., Bechmann, A. & Petersen, M.B. (2020). 'Partisan Polarization is the Primary Psychological Motivation Behind "Fake News" Sharing on Twitter'. 25 March 2020, https://psyarxiv.com/v45bk/

Pakistani Prevention of Electronic Crimes Act of 2016. http://www.na.gov.pk/uploads/documents/1470910659_707.pdf

Pamment, J., Nothaft, H., Agardh-Twetman, H. & Fjällhed. (2018). 'Countering Information Influence Activities: A handbook for communicators'. https://www.msb.se/RibData/Filer/pdf/28698.pdf

Paquette, D. (2019). 'Nigeria's 'fake news' bill could jail people for lying on social media. Critics call it censorship.' *The Washington Post, 25* November 2019. https://www.washingtonpost.com/world/africa/nigerias-fake-news-bill-could-jail-people-for-lying-on-social-media-critics-call-it-censorship/2019/11/25/ccf33c54-0f81-11ea-a533-90a7becf7713_story.html

Palma, S., Munshi, N. & Reed, J. (2020). 'Singapore 'falsehoods' law shows perils of fake news fight'. *Financial Times*, 04 February 2020. https://www.ft.com/content/e50eb042-3db3-11ea-a01a-bae547046735

Parkinson, H. J. (2016). 'Click and Elect: How Fake News Helped Donald Trump Win a Real Election'. *The Guardian*, 14 November 2016. https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies

Partnership on AI. (2020). 'The Deepfake Detection Challenge: Insights and Recommendations for AI and Media Integrity'. 12 March 2020. https://www.partnershiponai.org/a-report-on-the-deepfake-detection-challenge/

Pasternack, A. (2020). 'Facebook is quietly pressuring its independent fact-checkers to change their rulings'. *Fast Company*, 20 August 2020. https://www.fastcompany.com/90538655/facebook-is-quietly-pressuring-its-independent-fact-checkers-to-change-their-rulings

Paul, R. (2018). 'Bangladesh shuts down high-speed mobile internet on election eve'. *Reuters*, 29 December 2018. https://www.reuters.com/article/us-bangladesh-election-internet/bangladesh-shuts-down-high-speed-mobile-internet-on-election-eve-idUSKCN1OS0AR

Penney, V. (2020). 'How Facebook Handles Climate Disinformation'. *New York Times*, 14 July 2020. https://www.nytimes.com/2020/07/14/climate/climate-facebook-fact-checking.html

Pennycook, G., Bear, A. Collins, E. & Rand, D.G. (2020) "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings." Working paper, Last revised: 16 Jan 2020. http://dx.doi.org/10.2139/ssrn.3035384

Pennycook, G. & Rand, D.G. (2019). 'Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning.' *Cognition* 188, July 2019. pp39-50. https://doi.org/10.1016/j.cognition.2018.06.011

People's Daily. (2013). '国家互联网信息办部署打击网络谣言'. 03 May 2013. http://politics.people.com.cn/n/2013/0503/c1001-21348755.html

Perraudin, F. (2019). 'Twitter accuses Tories of misleading public with 'factcheck' foray'. *The Guardian*, 20 November 2019. https://www.theguardian.com/politics/2019/nov/20/twitter-accuses-tories-of-misleading-public-in-factcheck-row

Philippines Revised Penal Code, Act no 10951. (2017). https://www.officialgazette.gov.ph/downloads/2017/08aug/20170829-RA-10951-RRD.pdf

Philippines Senate. (2019). Bill no 9, Anti-False Content Bill. An Act prohibiting the publication and proliferation of false content on the Philippine internet, providing measures to counteract its effects and prescribing penalties therefore. https://senate.gov.ph/lisdata/3022527054!.pdf

Pichai, S. (2020). 'COVID-19: How we're continuing to help'. *Google*, 15 March 2020. https://blog.google/inside-google/company-announcements/covid-19-how-were-continuing-to-help/

PKULaw. (2017). 'Criminal Law of the People's Republic of China (2017 Amendment PKULAW Version) [Effective]'. 04 November 2017. http://en.pkulaw.cn/display.aspx?cgid=703dba7964330b85bdfb&lib=law

Plasilova, I., Hill, J., Carlberg, M., Goubet, M. & Procee, R. (2020). 'STUDY FOR THE "Assessment of the implementation of the Code of Practice on Disinformation". European Commission, May 2020. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=66649

PolitiFact (2018). 'Here's our statement on a recent story in The Guardian about third-party fact-checking on Facebook. Four other fact-checking groups signed as well'. *Twitter*, 13 December 2018. https://twitter.com/politifact/status/1073309698894585856?s=11

Pollock, S. (2019). 'Facebook fined €2m for under-reporting its complaints figures'. *The Times*, 07 July 2019. https://www.thetimes.co.uk/article/facebook-fined-2m-for-under-reporting-its-complaints-figures-7fhg26dct

Polonski, V. (2016). 'Impact of social media on the outcome of the EU referendum'. In EU Referendum Analysis 2016: Media, Voters and the Campaign. Jackson, D, Thorsen, E, Wring, D. Loughborough University Center for the Study of Journalism, Culture and Community. http://www.referendumanalysis.eu/

Pomares, J. & Guzman, N. (2015). 'The hardest check: Measuring the impact of fact-checking'. https://www.poynter.org/wp-content/uploads/2015/10/The-hardest-check-1.pdf

Posetti, J. (2013). 'The 'Twitterisation' of investigative journalism' in S. Tanner & N. Richardson (Eds.), Journalism Research and Investigation in a Digital World (pp. 88-100): *Oxford University Press*, Melbourne. https://ro.uow.edu.au/cgi/viewcontent.cgi?article=2765&context=lhapapers

Posetti, J. (2017). 'Fighting Back Against Prolific Online Harassment: Maria Ressa'. In '*An Attack on One is an Attack on All*. Successful Initiatives To Protect Journalists and Combat Impunity', Larry Kilman, UNESCO. https://en.unesco.org/sites/default/files/an_attack_on_on_is_an_attack_on_all_chapter_8.pdf

Posetti, J. (2017b). 'Protecting Journalism Sources in the Digital Age'. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000248054

Posetti, J. (2018a). 'Combatting online abuse: when journalists and their sources are targeted'. https://en.unesco.org/sites/default/files/module_7.pdf

Posetti, J. (2018b). ''News industry transformation: digital technology, social platforms and the spread of misinformation and disinformation' in Ireton & Posetti (eds) Journalism, 'Fake News' & Disinformation. (UNESCO) https://en.unesco.org/fightfakenews

Posetti, J. (2018c). 'Time to step away from the 'bright, shiny things'? Towards a sustainable model of journalism innovation in an era of perpetual change'. *Journalism Innovation Project, Reuters & the University of Oxford*. November 2018. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-11/Posetti_Towards_a_Sustainable_model_of_Journalism_FINAL.pdf

Posetti, J. (2020). 'Journalists like Maria Ressa face death threats and jail for doing their jobs. Facebook must take its share of the blame'. *CNN*, 02 July 2020. https://edition.cnn.com/2020/06/30/opinions/maria-ressa-facebook-intl-hnk/index.html

Posetti, J. & Bontcheva, K. (2020a). 'Disinfodemic – Deciphering COVID-19 disinformation'. Policy Brief 1. *UNESCO*, April 2020. https://en.unesco.org/sites/default/files/disinfodemic_deciphering_covid19_disinformation.pdf

Posetti, J. & Bontcheva, K. (2020b). 'Disinfodemic – Dissecting responses to COVID-19 disinformation'. Policy Brief 2. *UNESCO*, April 2020. https://en.unesco.org/sites/default/files/disinfodemic_deciphering_covid19_disinformation.pdf

Posetti, J. & Matthews, A. (2018). 'A Short Guide to the History of 'Fake News' and Disinformation. A Learning Module for Journalists and Journalism Educators'. *International Center For Journalists*, 23 July 2018. https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf

Posetti, J., Simon, F. & Shabbir, N. (2019a). 'Lessons in Innovation: How International News Organisations Combat Disinformation through Mission-Driven Journalism'. *Reuters*, April 2019. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-04/Posetti_Lessons_in_Innovation_FINAL.pdf

Posetti, J., Simon, F. & Shabbir, N. (2019b). 'What if Scale Breaks Community? Rebooting Audience Engagement When Journalism is Under Fire'. *Reuters*, October 2019. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-10/Posetti%20What%20if%20FINAL.pdf

Powell, A. (2020). 'Washington Post's Baron sends along the Class of 2020 with message that facts, truth matter'. *The Harvard Gazette*, 28 May 2020. https://news.harvard.edu/gazette/story/2020/05/martin-barons-message-to-class-of-2020-facts-and-truth-matter/?utm_medium=socialmedia&utm_source=hkstwitter

Power, J. (2020). 'Coronavirus vaccine: anti-vax movement threatens Asian recovery'. *SCMP, The Coronvirus Pandemic*, 01 June 2020. https://www.scmp.com/week-asia/health-environment/article/3086761/coronavirus-vaccine-anti-vax-movement-threatens-asian

Poynter. 'Commit to transparency — sign up for the International Fact-Checking Network's code of principles'. *IFCN*, *Poynter*. https://ifcncodeofprinciples.poynter.org/

Prager, A. (2019). 'Germany wants to fine Facebook over hate speech reporting'. *EurActiv*, 03 July 2019. https://www.euractiv.com/section/digital/news/germany-wants-to-fine-facebook-over-online-hate-speech-reporting/

Privacy International. (2020). 'No, Facebook is not telling you everything'. 24 February 2020. https://privacyinternational.org/long-read/3372/no-facebook-not-telling-you-everything

Proctor, K. (2020). 'UK anti-fake news unit dealing with up to 10 false coronavirus articles a day'. 30 March 2020. https://www.theguardian.com/world/2020/mar/30/uk-anti-fake-news-unit-coronavirus

Public Media Alliance. (2020). 'Emergency funds for US public broadcasting'. 01 April 2020. https://www.publicmediaalliance.org/emergency-funds-for-us-public-broadcasting/

Quadir, S. (2018). 'Tough New Bangladesh Measure Becomes Law, Seen Curbing Free Speech', *Reuters*, 28 October 2018. https://www.reuters.com/article/us-bangladesh-politics-journalism/tough-new-bangladesh-measure-becomes-law-seen-curbing-free-speech-idUSKCN1MI0NE

Quattrociocchi, W., Scala, A. & Sunstein, C. R. (2016). 'Echo Chambers on Facebook'. *Social Science Research Network*, 15 June 2016. https://papers.ssrn.com/abstract=2795110

Qui, L. (2020). 'Analyzing the Patterns in Trump's Falsehoods About Coronavirus'. *The New York Times*, 27 March 2020. https://www.nytimes.com/2020/03/27/us/politics/trump-coronavirus-factcheck.html

Qui, S. & Woo, R. (2018). 'China launches platform to stamp out 'online rumors'. *Reuters*, 30 August 2018. https://www.reuters.com/article/us-china-internet/china-launches-platform-to-stamp-out-online-rumors-idUSKCN1LF0HL

Quinn, C. (2020). 'Hungary's Orban Given Power to Rule By Decree With No End Date'. *Foreign Policy*, 31 March 2020. https://foreignpolicy.com/2020/03/31/hungarys-orban-given-power-to-rule-by-decree-with-no-end-date/

Rahman, S., Tully, P. & Foster, L. (2019). 'Attention is All They Need: Combatting Social Media Information Operations With Neural Language Models'. *Fire Eye*, 14 November 2019. https://www.fireeye.com/blog/threat-research/2019/11/combatting-social-media-information-operations-neural-language-models.html

Ranking Digital Rights. (2019). '2019 RDR Corporate Accountability Index'. https://rankingdigitalrights.org/index2019/

Ranking Digital Rights. (2020). 'RDR Corporate Accountability Index: Transparency and accountability standards for targeted advertising and algorithmic systems: Pilot Study and Lessons Learned', 16 March 2020. https://rankingdigitalrights.org/wp-content/uploads/2020/03/pilot-report-2020.pdf

Rappler Research Team. (2018). 'Tip of the iceberg: Tracing the network of spammy pages in Facebook takedown'. 27 October 2018. https://www.rappler.com/newsbreak/investigative/215256-tracing-spammy-pages-network-facebook-takedown

Rauchfleisch, A. & Kaiser, J. (2020). 'The False Positive Problem of Automatic Bot Detection in Social Science Research'. *Berkman Klein Center Research Publication No. 2020-3*, March 2020. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565233

Read, M. (2016). 'Donald Trump won because of Facebook'. *New York Magazine*, 09 November 2016. https://nymag.com/intelligencer/2016/11/donald-trump-won-because-of-facebook.html

Reda, J. (2017). 'When filters fail: These cases show we can't trust algorithms to clean up the internet'. 28 September 2017. https://juliareda.eu/2017/09/when-filters-fail/

Redação Exame. (2019). '*Após derrubar veto de Bolsonaro, Congresso instaura CPI das Fake News, EXAME*'. 2 September 2019 https://exame.abril.com.br/brasil/apos-derrubar-veto-de-bolsonaro-congresso-instaura-cpi-das-fake-news/

Reid, A. & Dotto. C. (2019). 'Thousands of misleading Conservative ads side-step scrutiny thanks to Facebook policy'. *First Draft*, 06 December 2019. https://firstdraftnews.org/latest/thousands-of-misleading-conservative-ads-side-step-scrutiny-thanks-to-facebook-policy/

Reppell, L. & Shein, E. (2019). 'Disinformation Campaigns and Hate Speech: Exploring the Relationship and Programming Interventions'. *IFES*, 26 April 2019. https://www.ifes.org/publications/disinformation-campaigns-and-hate-speech-exploring-relationship-and-programming

Repnikova, M. (2018). 'China's Lessons for Fighting Fake News'. *FP*, 06 September 2018. https://foreignpolicy.com/2018/09/06/chinas-lessons-for-fighting-fake-news/

Resnick, P. Ovadya, A. & Gilchrist, G. (2019). 'Iffy Quotient: A Platform Health Metric for Misinformation'. *Social Media Responsibility*, *University of Michigan*, 23 July 2019. https://csmr.umich.edu/wp-content/uploads/2019/07/UMSI-CSMR-Iffy-Quotient-Whitepaper-v2.pdf

Ressa, M. A. (2016). 'Propaganda war: Weaponizing the internet'. *Rappler*, 7 February 2019. https://www.rappler.com/nation/148007-propaganda-war-weaponizing-internet

Ressa, M. (2019). 'When journalists are under attack, democracy is under attack'. *Daily Maverick*, 30 September 2019. https://www.dailymaverick.co.za/article/2019-09-30-when-journalists-are-under-attack-democracy-is-under-attack/

Revelli, A. & Foster, L. (2020). '"Distinguished Impersonator" Information Operation That Previously Impersonated U.S. Politicians and Journalists on Social Media Leverages Fabricated U.S. Liberal Personas to Promote Iranian Interests'. *FireEye*, 12 February 2020. https://www.fireeye.com/blog/threat-research/2020/02/information-operations-fabricated-personas-to-promote-iranian-interests.html

RFE/RL. (2018a). 'Belarus Passes Legislation Against 'Fake News' Media'. 14 June 2018. https://www.rferl.org/a/belarus-assembly-passes-controversial-fake-news-media-legislation/29291033.html

RFE/RL. (2018b). 'Kazakhstan Shuts Down Independent News Site'. 28 May 2018. https://www.rferl.org/a/kazakhstan-shuts-down-independent-news-site-ratel/29254964.html

Richter, A. (2019) 'Disinformation in the media under Russian law'. *IRIS Extra, European Audiovisual Observatory*, June 2019. https://rm.coe.int/disinformation-in-the-media-under-russian-law/1680967369

Robinson, O., Coleman, A., Sardarizadeh, S. (2019). 'A Report of Anti-Disinformation Initiatives'. *OXTEC*, August 2019. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/08/A-Report-of-Anti-Disinformation-Initiatives

Rodriguez-Ferrand, G. (2019). 'Initiatives to Counter Fake News: Argentina'. *Library of Congress*, April 2019. https://www.loc.gov/law/help/fake-news/argentina.php#_ftnref15

Romm, T., Dwoskin, E. & Timberg, C. (2019). 'Sri Lanka's social media shutdown illustrates global discontent with Silicon Valley'. *The Washington Post*, 22 April 2019. https://www.washingtonpost.com/technology/2019/04/22/sri-lankas-social-media-shutdown-illustrates-global-discontent-with-silicon-valley/

Romm, T. Stanley-Becker, I. & Timberg, C. (2020). '"Facebook won't limit political ad targetin or stop false claims under new ad rules'. *The Washington Post*, 09 January 2020. https://www.washingtonpost.com/technology/2020/01/09/facebook-wont-limit-political-ad-targeting-or-stop-pols-lying/

Rosen, G. (2020). 'Investments to Fight Polarization'. *Facebook Newsroom*, 27 May 2020. https://about.fb.com/news/2020/05/investments-to-fight-polarization/

Rosen, G. & Lyons, T. (2019). 'Remove, Reduce, Inform: New Steps to Manage Problematic Content'. *Facebook Newsroom*, 10 April 2019. https://about.fb.com/news/2019/04/remove-reduce-inform-new-steps/

Rosling, H. (2018). 'Factfulness: Ten Reasons We're Wrong About The World–And Why Things Are Better Than You Think'. Sceptre.

Rössler, A., Cozzolino, D., Veroliva, L., Riess, C. Thies, J. & Neiβner, M. (2018). 'FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces'. 24 March 2018. https://arxiv.org/abs/1803.09179

Roth, Y. (2019). 'Information operations on Twitter: principles, process, and disclosure'. *Twitter Blog*, 13 June 2019. https://blog.twitter.com/en_us/topics/company/2019/information-ops-on-twitter.html

Roth, Y. & Achuthan, A. (2020). 'Building rules in public: Our approach to synthetic & manipulated media'. *Twitter Blog*, 04 February 2020. https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media.html

Roucaute, D. (2017). '« Le Monde » s'engage dans l'éducation à l'information'. *Le Monde*, 02 February, 2017 (updated 27 November 2017). https://www.lemonde.fr/les-decodeurs/article/2017/02/02/le-monde-s-engage-dans-l-education-a-l-information_5073215_4355770.html

RSF. (2018). 'Online Harassment of Journalists - Attack of the trolls'. https://rsf.org/sites/default/files/rsf_report_on_online_harassment.pdf

RSF (2019). 'Burkina Faso : l'amendement du code pénal doit être déclaré inconstitutionnel'. July 17, 2019. https://rsf.org/fr/actualites/burkina-faso-lamendement-du-code-penal-doit-etre-declare-inconstitutionnel-2

Ruan, L., Knockel, J., Q. Ng, J., & Crete-Nishihata, M. (2016). 'One App, Two Systems: How WeChat uses one censorship policy in China and another internationally'. *The Citizen Lab*, 30 November 2016. https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems/

Sabbagh, D. (2020). 'Cross-Whitehall unit set up to counter false coronavirus claims'. *The Guardian*, 09 March 2020. https://www.theguardian.com/world/2020/mar/09/cross-whitehall-unit-coronavirus-disinformation

Sadek, G. (2019). 'Initiatives to Counter Fake News: Egypt'. *Library of Congress*, April 2019. https://www.loc.gov/law/help/fake-news/egypt.php#_ftn11

Sally Chan, M., Jones, C. R., Hall Jamieson, K. & Albarracin, D. (2017). 'Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation'. *Psychological Science*, 12 September 2017. https://journals.sagepub.com/doi/full/10.1177/0956797617714579

Sanders, E. (2020). 'Washington State Sues Facebook for "Repeatedly" Violating Campaign Finance Law'. *The Stranger*, 14 April 2020. https://www.thestranger.com/slog/2020/04/14/43407935/washington-state-sues-facebook-for-repeatedly-violating-campaign-finance-law

Sang-Hun, C. (2018). 'South Korea Declares War on 'Fake News', Worrying Government Critics'. *The New York Times*. 02 October 2018. https://www.nytimes.com/2018/10/02/world/asia/south-korea-fake-news.html

Sankin, A. (2020). 'Want to Find a Misinformed Public? Facebook's Already Done It'. *The Markup*, 23 April 2020. https://themarkup.org/coronavirus/2020/04/23/want-to-find-a-misinformed-public-facebooks-already-done-it

Satariano, A. (2018). 'Ireland's Abortion Referendum Becomes a Test for Facebook and Google', *New York Times*, 25 May 2018. https://www.nytimes.com/2018/05/25/technology/ireland-abortion-vote-facebook-google.html

Satariano, A. & Tsang, A. (2019). 'Who's Spreading Disinformation in U.K. Election? You Might Be Surprised'. *New York Times*, 10 December 2019. https://www.nytimes.com/2019/12/10/world/europe/elections-disinformation-social-media.html

Schiffrin, A. & Goodman, E. (2019). 'AI Startups and the Fight Against Online Disinformation'. *The German Marshall Fund of the United States*, 05 September 2019. http://www.gmfus.org/publications/ai-startups-and-fight-against-online-disinformation

Schmerling, R. H. (2020). 'Be careful where you get your news about coronavirus'. *Harvard Health Publishing*, 01 February 2020. https://www.health.harvard.edu/blog/be-careful-where-you-get-your-news-about-coronavirus-2020020118801

Schmitt-Beck, R. (2008). 'Bandwagon effect'. In W. Donsbach (Ed.) 'The international encyclopedia of communication', Vol. 2 (pp. 308–310). Oxford, UK. Wiley-Blackwell

Schroepfer, M. (2019). 'Creating a data set and a challenge for deepfakes'. *Facebook AI*, 05 September 2019. https://ai.facebook.com/blog/deepfake-detection-challenge/

Schulman, S. (2019). 'Development, digitalization and disinformation in Myanmar'. *U brief*. May 2019. https://www.ui.se/globalassets/ui.se-eng/publications/ui-publications/2019/ui-brief-no.-5-2019.pdf

Schulten, K. & Brown, A. C. (2017). 'Evaluating Sources in a 'Post-Truth' World: Ideas for Teaching and Learning About Fake News'. New York Times, 19 January 2017. https://www.nytimes.com/2017/01/19/learning/lesson-plans/evaluating-sources-in-a-post-truth-world-ideas-for-teaching-and-learning-about-fake-news.html

Schumaker, E. (2019). 'Why Conspiracy Theories Work so Well on Facebook'. OneZero, 05 March 2019. https://onezero.medium.com/why-conspiracy-theories-work-so-well-on-facebook-466a42af6b76

Scott, C. (2018). 'With an interactive game, the BBC is helping young people better understand the disinformation ecosystem'. Journalism.co.uk, 01 May 2018. https://www.journalism.co.uk/news/with-an-interactive-game-the-bbc-is-helping-young-people-better-understand-the-disinformation-ecosystem-/s2/a721168/

See Kit, T. (2019). 'Law to curb deliberate online falsehoods takes effect'. *CNA*. 02 October 2019. https://www.channelnewsasia.com/news/singapore/law-to-curb-deliberate-online-falsehoods-takes-effect-11962068

Sen, A. & Zadrozny, B. (2020). 'QAnon groups have millions of members on Facebook, documents show'. *NBC News*, 10 August 2020. https://www.nbcnews.com/tech/tech-news/qanon-groups-have-millions-members-facebook-documents-show-n1236317

Sénécat, A. (2018). 'Les fausses informations circulent de moins en moins sur Facebook'. *Le Monde*, 17 October 2018. https://www.lemonde.fr/les-decodeurs/article/2018/10/17/les-fausses-informations-perdent-du-terrain-sur-facebook_5370461_4355770.html

Shin, J. & Thorsen, K. (2017). 'Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media'. *Journal of Communication*, 28 February 2017. https://academic.oup.com/joc/article/67/2/233/4082394

Shu, C. & Shieber, J. (2020). 'Facebook, Reddit, Google, LinkedIn, Microsoft, Twitter and YouTube issue joint statement on misinformation'. *TechCrunch*, 17 March 2020. https://techcrunch.com/2020/03/16/facebook-reddit-google-linkedin-microsoft-twitter-and-youtube-issue-joint-statement-on-misinformation/

Silva, M., Oliveira, L. S., Andreou, A., Vaz de Melo, P. O., Goga, O. & Benevenuto, F. (2020). 'Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook'. In *Proceedings of the Web Conference* (*WWW'20*), April 2020. https://doi.org/10.1145/3366423.3380109

Silverman, C. et al. (2014). 'The Verification Handbook (European Journalism Centre). https://verificationhandbook.com/book/index.php

Silverman, C. (2016). 'Fake News Expert On How False Stories Spread And Why People Believe Them'. Fresh Air – transcript on *NPR*, 14 December 2016. https://www.npr.org/2016/12/14/505547295/fake-news-expert-on-how-false-stories-spread-and-why-people-believe-them?t=1589899554744

Silverman, C. (2017a). '5 ways scammers exploit Facebook to feed you false information'. *BuzzFeed*, 28 April 2017. https://www.buzzfeed.com/craigsilverman/how-facebook-is-getting-played

Silverman, C. (2017b). 'Facebook says its Fact-Checking Program helps reduce the spread of a fake story by 80%'. *Buzzfeed*, 11 October 2017. https://www.buzzfeednews.com/article/craigsilverman/facebook-just-shared-the-first-data-about-how-effective-its

Silverman, C. (2019). 'How A Massive Facebook Scam Siphoned Millions Of Dollars From Unsuspecting Boomers'. *Buzzfeed*, 16 October 2019. https://www.buzzfeednews.com/article/craigsilverman/facebook-subscription-trap-free-trial-scam-ads-inc

Silverman, C. (Ed.) (2020). 'Verification Handbook. For Disinformation And Media Manipulation'. *European Journalism Centre*. https://datajournalism.com/read/handbook/verification-3

Silverman, C. & Alexander, L. (2016). 'How Teens In The Balkans Are Duping Trump Supporters With Fake News'. *BuzzFeed*, 3 November 2016. https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo

Silverman, C., Lytvynenko, J. & Kung, W. (2020). 'Disinformation For Hire: How a New Breed of PR Firms Is Selling Lies Online'. *BuzzFeed*, 6 January 2020. https://www.buzzfeednews.com/article/craigsilverman/disinformation-for-hire-black-pr-firms

Silverman, C., Lytvynenko, J. & Pham, S. (2017). 'These Are 50 Of The Biggest Fake News Hits On Facebook In 2017'. *Buzzfeed*, 28 December 2017. https://www.buzzfeednews.com/article/craigsilverman/these-are-50-of-the-biggest-fake-news-hits-on-facebook-in

Silverman, C. & Pham, S. (2018). '*These Are 50 Of The Biggest Fake News Hits On Facebook In 2018*'. *Buzzfeed*, 28 December 2018. https://www.buzzfeednews.com/article/craigsilverman/facebook-fake-news-hits-2018

Simon, J. (2020). 'COVID-19 is spawning a global press-freedom crackdown'. *CJR*, 25 March 2020. https://www.cjr.org/analysis/coronavirus-press-freedom-crackdown.php

Singapore Democratic Party (2020). 'SDP Files Case Against MOM In High Court To Fight For What Little Democratic Space We Have Left In S'pore'. 08 January 2020. https://yoursdp.org/news/sdp-files-case-against-mom-in-high-court-to-fight-for-what-little-democratic-space-we-have-left-in-s%27pore

Singapore POFMA (2019). 'Protection From Online Falsehoods and Manipulation Act 2019'. Singapore Statutes online. https://sso.agc.gov.sg/Acts-Supp/18-2019/Published/20190625?DocDate=20190625#pr11-

Sippitt, A. (2020). 'Playing the long game: the impact of fact checkers on public figures, institutions, and the media'. *Fullfact*, 13 March 2020. https://fullfact.org/blog/2020/mar/long-game-impact-fact-checkers/

Smith, B. (2020a). 'When Facebook Is More Trustworthy Than the President'. *New York Times*, 15 March 2020. https://www.nytimes.com/2020/03/15/business/media/coronavirus-facebook-twitter-social-media.html

Smith, A. (2020b). 'Zuckerberg Criticises Twitter After It Fact-Checks Trump Tweets, Saying It Shouldn't Be 'Arbiter Of Truth''. *Independent*, 28 May 2020. https://www.independent.co.uk/life-style/gadgets-and-tech/news/trump-twitter-fact-check-mark-zuckerberg-jack-dorsey-a9536051.html?amp

Smith, A. & Shabad, R. (2020). 'Trump signs executive order aimed at social media companies after fuming over fact-check'. *NBC News*, 28 May 2020. https://www.nbcnews.com/politics/white-house/angry-over-how-social-media-platforms-are-treating-him-trump-n1216401

Smyrnaios, N., Chauvet, S. & Marty, E. (2017). 'The Impact of CrossCheck on Journalists & the Audience.' *First Draft*. https://firstdraftnews.org/wp-content/uploads/2017/11/Crosscheck_rapport_EN_1129.pdf/

Snapchat. (2017). 'Introducing the new Snapchat'. *Snap Inc*, 29 November 2017. https://www.snap.com/en-US/news/post/introducing-the-new-snapchat/

Snapchat. (2020). 'Coronavirus: How Snapchatters are reacting to the news and staying informed'. 17 March 2020. https://forbusiness.snapchat.com/blog/coronavirus-how-snapchatters-are-reacting-to-the-news-and-staying-informed

Soares, I. (2017) 'The Macedonia Story'. *CNN*. https://money.cnn.com/interactive/media/the-macedonia-story/

Solaiman, I, Clark, J. & Brundage, M. (2019). 'GPT-2: 1.5B Release'. *Open AI*, 5 November 2019. https://openai.com/blog/gpt-2-1-5b-release/

Soltani, A. (2018). 'Oral evidence: Disinformation and 'fake news'. HC 363. Questions 4274 – 4382. Published on 27 November 2018. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture- media-and-sport-committee/disinformation-and-fake-news/oral/92924.html

Soundararajan, T., Kumar, A., Nair, P., Greely, J. (2019). 'Facebook India Towards the Tipping Point of Violence: Caste And Religious Hate Speech'. *Equality Labs*. https://www.equalitylabs.org/facebookindiareport

South African Government. (2020). Amendments to the 'Disaster Management Act, 2002, 18 March 2020. https://www.gov.za/sites/default/files/gcis_document/202003/43107gon318.pdf

Sovuthy, K. (2018). 'Government to tackle fake news frenzy'. *Khmer Times*, 05 July 2018. https://www.khmertimeskh.com/508265/government-to-tackle-fake-news-frenzy/

Spencer, S. (2019). 'An update on our political ads policy'. *Google Ads*, 20 November 2019. https://www.blog.google/technology/ads/update-our-political-ads-policy/

Spiegel, E. (2017). 'How Snapchat is separating social from media'. *Axios*, 29 November 2017. https://www.axios.com/how-snapchat-is-separating-social-from-media-2513315946.html

Spinney, L. (2019). 'Fighting Ebola is hard. In Congo, fake news makes it harder'. *AAAS Science Mag*, 14 January 2019. https://www.sciencemag.org/news/2019/01/fighting-ebola-hard-congo-fake-news-makes-it-harder

Spivak, C. (2010). 'The Fact-Checking Explosion. American Journalism Review'. University of Maryland. Winter 2010

Spring, M. (2020). 'Coronavirus: The seven types of people who start and spread viral misinformation'. *BBC News*, 04 May 2020. https://www.bbc.co.uk/news/blogs-trending-52474347?utm_source=First+Draft+Subscribers&utm_campaign=41894fb736-EMAIL_CAMPAIGN_2020_03_26_08_17_COPY_01&utm_medium=email&utm_term=0_2f24949eb0-41894fb736-265419209&mc_cid=41894fb736&mc_eid=a7e431ef92

Sri Lanka Brief. (2019). 'Sri Lanka govt to bring new laws to curb fake news & hate speech that jeopardize national harmony & national security'. 06 June 2019. https://srilankabrief.org/2019/06/sri-lanka-govt-to-bring-new-laws-to-curb-fake-news-hate-speech-that-jeopardize-national-harmony-national-security/

Srnicek, N. (2017). 'Platform Capitalism'. John Wiley & Sons, U.S

Stanley-Becker, I. (2020) 'Google Greenlights ads with "blatant disinformation" about voting by mail'. *The Washington Post*, 28 August 2020. https://www.washingtonpost.com/technology/2020/08/28/google-ads-mail-voting/?utm_source=twitter&utm_medium=social&utm_campaign=wp_main

Stanton, D. (2019). 'Advancing research on fake audio detection'. *Google News Initiative*, 31 January 2019. https://blog.google/outreach-initiatives/google-news-initiative/advancing-research-fake-audio-detection/

Staff, T. (2019). 'Election judge bars anonymous internet ads despite Likud objection'. *The Times of Israel*, 28 February, 2019. https://www.timesofisrael.com/election-judge-bars-anonymous-internet-adds-despite-likud-objection/

Starbird, K. (2017). 'Information Wars: A Window into the Alternative Media Ecosystem'. *In Proceedings of the 11th International AAAI Conference on Web and Social Media* (*ICWSM-17*), 15 March 2017. https://medium.com/hci-design-at-uw/information-wars-a-window-into-the-alternative-media-ecosystem-a1347f32fd8f

Statt, N. (2020). 'YouTube is a $15 billion-a-year business, Google reveals for the first time'. *The Verge*, 03 February 2020. https://www.theverge.com/2020/2/3/21121207/youtube-google-alphabet-earnings-revenue-first-time-reveal-q4-2019

Stecklow, S. (2018). 'Why Facebook is losing the war on hate speech in Myanmar'. *Reuters*, 15 August 2018. https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/

Stewart, E. (2019). 'Facebook is refusing to take down a Trump ad making false claims about Joe Biden'. *Vox*, 9 October 2019. https://www.vox.com/policy-and-politics/2019/10/9/20906612/trump-campaign-ad-joe-biden-ukraine-facebook

Stolton, S. (2020) 'EU Rapid Alert System used amid coronavirus disinformation campaign'. EurActiv, 04 March 2020. https://www.euractiv.com/section/digital/news/eu-alert-triggered-after-coronavirus-disinformation-campaign/

Storm, H. (2020). 'Media ethics, safety and mental health: reporting in the time of Covid-19'. *EJN*, 12 March 2020. https://ethicaljournalismnetwork.org/media-ethics-safety-and-mental-health-reporting-in-the-time-of-covid-19

Strobelt, H. & Gehrmann, S. (2019). 'Catching a Unicorn with GLTR: A tool to detect automatically generated text'. http://gltr.io/

Suárez, E. (2020). '"I think removing content by Donald Trump will always have to meet a very high bar". *Reuters Institute for the study of Journalism*, *University of Oxford*, 19 May 2020. https://reutersinstitute.politics.ox.ac.uk/risj-review/i-think-removing-content-donald-trump-will-always-have-meet-very-high-bar

Subrahmanian, V.S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A., & Menczer, F. (2016). 'The DARPA Twitter Bot Challenge'. *Computer* Volume 49 Issue 6: pp.38–46. https://doi.org/10.1109/MC.2016.183

Swart, J., Peters, C., & Broersma, M. (2017). 'Navigating cross-media news use: Media repertoires and the value of news in everyday life'. *Journalism Studies*, 18(11), 1343-1362. https://doi.org/10.1080/1461670X.2015.1129285

Swedish Government (2018). En ny myndighet för psykologiskt försvar. Dir 2018-80. http://www.sou.gov.se/wp-content/uploads/2019/05/Direktiv-2018-80-En-ny-myndighet-för-psykologiskt-försvar.pdf

Swire, B. & Ecker, U.K.H. (2018). 'Misinformation and its Correction: Cognitive Mechanisms and Recommendations for Mass Communication'. In Southwell, B, Thorson, E.A. & Sheble, L (Eds.), '*Misinformation and Mass Audiences*'. Austin: University of Texas Press.

Sydell, L. (2016). 'We Tracked Down A Fake-News Creator In The Suburbs. Here's What We Learned'. *NPR*, 23 November 2016. http://www.cpr.org/news/npr-story/npr-finds-the-head-of-a-covert-fake-news-operation-in-the-suburbs

Symonds, T. (2018). 'The Skripals and the Salisbury poisoning: What happened next?'. *BBC News*, 28 December 2018. https://www.bbc.com/news/uk-46538255

Szunyogh, B. (1955). 'Psychological warfare; an introduction to ideological propaganda and the techniques of psychological warfare'. *William-Frederick Press*, United States. p. 13. Retrieved 2015-02-11.

Taibbi, M. (2019). 'YouTube, Facebook Purges Are More Extensive Than You Think'. Rolling Stone, 07 June 2019. https://www.rollingstone.com/politics/politics-features/youtube-facebook-purges-journalists-845790/

Tameez, H. (2020). 'YouTube's algorithm is pushing climate misinformation videos, and their creators are profiting from it'. *Nieman Lab*, 16 January 2020. https://www.niemanlab.org/2020/01/youtubes-algorithm-is-pushing-climate-misinformation-videos-and-their-creators-are-profiting-from-it/

Tanakasempipat, P. (2019a). 'Thailand passes internet security law decried as 'cyber martial law'. *Reuters*. 28 February 2019. https://www.reuters.com/article/us-thailand-cyber/thailand-passes-internet-security-law-decried-as-cyber-martial-law-idUSKCN1QH1OB

Tanakasempipat, P. (2019b). 'Thailand unveils 'anti-fake news' center to police the internet'. *Reuters*, 1 November 2019. https://www.reuters.com/article/us-thailand-fakenews/thailand-unveils-anti-fake-news-center-to-police-the-internet-idUSKBN1XB48O

Tapsell, R. (2019). 'Arrests for political hoax news in Indonesia'. *East Asia Forum*. 08 November 2019. https://www.eastasiaforum.org/2019/11/08/arrests-for-political-hoax-news-in-indonesia/

Tarbush, B., Teytelboym, A. (2012). 'Homophily in Online Social Networks'. *In*: *Goldberg P.W.* (*eds*) *Internet and Network Economics*. WINE 2012. Lecture Notes in Computer Science, vol 7695. Springer. https://doi.org/10.1007/978-3-642-35311-6_40

Tardáguila, C. (2019). 'Fact-checkers launch Instagram, WhatsApp and Telegram stickers to gently warn about false news'. *Poynter*, 05 June 2019. https://www.poynter.org/fact-checking/2019/fact-checkers-launch-instagram-whatsapp-and-telegram-stickers-to-gently-warn-about-false-news/

Tardáguila, C. (2020). 'Coronavirus: Fact-checkers from 30 countries are fighting 3 waves of misinformation'. *Poynter*, 28 January 2020. https://www.poynter.org/fact-checking/2020/coronavirus-fact-checkers-from-30-countries-are-fighting-3-waves-of-misinformation/

Tardáguila, C., Benevenuto, F. & Ortellado, P. (2018). 'Fake News Is Poisoning Brazilian Politics. WhatsApp Can Stop It.' *The New York Times*, 17 October 2018. https://www.nytimes.com/2018/10/17/opinion/brazil-election-fake-news-whatsapp.html

Taylor, L. (2020). 'The coronavirus story is unfathomably large. We must get the reporting right'. *The Guardian*, 21 March 2020. https://www.theguardian.com/media/2020/mar/22/the-coronavirus-story-is-unfathomably-large-we-must-get-the-reporting-right

@Team Trump. (2019). '@facebook wants to take important tools away from us for 2020.' *Twitter*, 20 November 2019. https://twitter.com/TeamTrump/status/1197217636662337536?s=20

Tennessee General Assembly Fiscal Review Committee. (2020). 'Fiscal Note HJR 779'. 04 February 2020. http://www.capitol.tn.gov/Bills/111/Fiscal/HJR0779.pdf

Tennessee Legislature. (2020). 'House Joint Resolution 779'. 03 October 2019. http://www.capitol.tn.gov/Bills/111/Bill/HJR0779.pdf

Teyssou, D., Leung, J.M., Apostolidis, E., Apostolidis, K., Papadopoulos, S., Zampoglou, M., & Mezaris, V. (2017). 'The InVID Plug-in: Web Video Verification on the Browser. *In Proceedings of the First International Workshop on Multimedia Verification*'. (*pp. 23-30*). *ACM*, October 2017. https://www.researchgate.net/publication/320570485_The_InVID_Plug-in_Web_Video_Verification_on_the_Browser

Thailand Computer Crime Act. (2017). https://en.wikisource.org/wiki/Translation:Computer_Crimes_Act_(No._2)_2017

Thamm, M. (2019). 'Navigating your way in a world filled with untruths'. *Daily Maverick*, 26 September 2019. https://www.dailymaverick.co.za/article/2019-09-26-navigating-your-way-in-a-world-filled-with-untruths/

Theisen, W., Brogan, J., Bilo Thomas, P., Moreira, D., Phoa, P., Weninger, T. & Scheirer, W. (2020). 'Automatic Discovery of Political Meme Genres with Diverse Appearances' *ArXiv e-print*, January 2020. https://ui.adsabs.harvard.edu/abs/2020arXiv200106122T/abstract

The Jakarta Post. (2016). 'ITE Law draft revision passed into law'. 27 October 2016. https://www.thejakartapost.com/news/2016/10/27/ite-law-draft-revision-passed-into-law.html

The Jakarta Post. (2019). 'Stop Hoax Indonesia program to educate internet users in 17 cities'. 11 August 2019. https://www.thejakartapost.com/life/2019/08/10/stop-hoax-indonesia-program-to-educate-internet-users-in-17-cities.html

Theocharis, Y., Barberà, P., Fazekas, Z., Popa, S. A.,& Parnet, O. (2016). 'A bad workman blames his tweets: the consequences of citizens' uncivil twitter use when interacting with party candidates'. Journal of Communication 66(6): 1007–1031, 28 October 2016. https://onlinelibrary.wiley.com/doi/full/10.1111/jcom.12259

Thomas, P. (2020). 'Coronavirus: Media pundits call for networks to stop airing Trump's 'misleading' briefings', *Independent*, 22 March 2020. https://www.independent.co.uk/news/world/americas/us-politics/coronavirus-trump-white-house-briefing-rachel-maddow-margaret-sullivan-jennifer-senior-a9416741.html

Thompson, S. (2020). 'Fact Check: As Of March 7, 2020, Tanzania And Zambia Had NOT Confirmed First Cases Of Coronavirus'. *Lead Stories*, 07 March 2020. https://leadstories.com/hoax-alert/2020/03/fact-check-tanzania-confirms-first-case-of-coronavirus---mcm.html

Thompson, T. (2019). 'Countering Russian disinformation the Baltic nations' way'. *The Conversation*, 9 January, 2019. https://theconversation.com/countering-russian-disinformation-the-baltic-nations-way-109366

Thorne, J., Vlachos, A., Christodoulopoulos, C., & Mittal, A. (2018). 'FEVER: a large-scale dataset for Fact Extraction and VERification'. Cornell University, 14 March 2018. https://arxiv.org/abs/1803.05355

Tidy, J. & Schraer, R. (2019). 'General election 2019: Ads are 'indecent, dishonest and untruthful'. *BBC*, 17 December 2019. https://www.bbc.co.uk/news/technology-50726500

TIMEP. (2019). 'TIMEP Brief: The Law Regulating the Press, Media, and the Supreme Council for Media Regulation'. 15 May, 2019. https://timep.org/reports-briefings/timep-brief-the-law-regulating-the-press-media-and-the-supreme-council-for-media-regulation/

TNN. (2020). 'How India Became The Global Leader in Internet Shutdowns'. *Times of India*, 08 January 2020. https://timesofindia.indiatimes.com/india/how-india-became-the-global-leader-in-internet-shutdowns/articleshow/72886376.cms?from=mdr

Tracy, M. (2020). 'News Media Outlets Have Been Ravaged by the Pandemic'. *The New York Times*, 10 May 2020. https://www.nytimes.com/2020/04/10/business/media/news-media-coronavirus-jobs.html

Trend Micro. (2020). 'Developing Story: COVID-19 Used in Malicious Campaigns'. 24 April 2020. https://www.trendmicro.com/vinfo/fr/security/news/cybercrime-and-digital-threats/coronavirus-used-in-spam-malware-file-names-and-malicious-domains

Troop, W. (2017). 'This Italian politician wants kids to become 'fake news hunters'. *The World*, 31 October 2017. https://www.pri.org/stories/2017-10-31/italian-politician-wants-kids-become-fake-news-hunters

Tsai, C.-H. & Brusilovsky, P. (2018). 'Beyond the Ranked List: User-Driven Exploration and Diversification of Social Recommendation'. *In Proceedings of the 23rd International Conference on Intelligent User Interfaces. ACM.* https://pennstate.pure.elsevier.com/en/publications/beyond-the-ranked-list-user-driven-exploration-and-diversificatio

Turvill, W. (2020). 'Going viral: Twitter and Facebook failing to contain Covid-19 misinformation'. *PressGazette*, 31 March 2020. https://www.pressgazette.co.uk/the-fake-news-epidemic-twitter-and-facebook-failing-to-contain-covid-19-misinformation/

Twitter. (2018). Oral Evidence for Lords Communications Committee – 'The Internet: To Regulate or Not To Regulate?' https://parliamentlive.tv/Event/Index/2cd62e7a-d3cf-4605-8d39-4fbaa0adaa76#player-tabs

Twitter. (2019a). 'New disclosures to our archive of state-backed information operations'. 20 December 2019. https://blog.twitter.com/en_us/topics/company/2019/new-disclosures-to-our-archive-of-state-backed-information-operations.html

Twitter. (2019b). 'Twitter Progress Report: Code of Practice against Disinformation'. https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019

UK Commons Select Committee on Digital, Culture, Media and Sport. (2018). *Parliamentarians from across the World to Question Richard Allan of Facebook, and the Information Commissioner at Inaugural Hearing of 'International Grand Committee' on Disinformation and 'Fake News'*, Press Release, 23 November 2018. https://www.parliament.uk/business/committees/committees-a-z/commons-select/digital-culture-media-and-sport-committee/news/grand-committee-evidence-17-19/

UK Commons Select Committee on Digital, Culture, Media and Sport. (2019). *Disinformation and 'Fake News': Final Report*. London: House of Commons. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

UK Department for Digital, Culture, Media and Sport (DCMS). (2018b). 'Oral evidence : Fake News – 08 February 2018' (George Washington University, Washington DC), HC 363. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culturemedia-and-sport-committee/fake-news/oral/78195.html

UK Department for Digital, Culture, Media and Sport (DCMS). (2018c). 'Disinformation and 'fake news': Interim Report: Government Response to the Committee's Fifth Report of Session 2017–19. HC 1630. Published on 23 October 2018. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1630/1630.pdf

UK Department for Digital, Culture, Media and Sport and UK Home Office (DCMS). (2019). 'Online Harms White Paper'. 12 February 2019. https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

UK Department for Digital, Culture, Media and Sport and UK Home Office (DCMS). (2020). 'Online Harms White Paper – Initial consultation response'. 12 February 2020. https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response

UK DCMS (Department for Digital, Culture, Media and Sport), UK Home Office, The Rt Hon Matt Hancock MP & The Rt Hon Sajid Javid MP. (2018a). 'New laws to make social media safer'. Press Release, 20 May 2018. https://www.gov.uk/government/news/new-laws-to-make-social-media-safer

UK DCMS. (Department for Digital, Culture, Media and Sport) & The Rt Hon Matt Hancock MP. (2018b). Policy Paper – 'Digital Charter', 25 January 2018. https://www.gov.uk/government/publications/digital-charter

UK Delegation to the OSCE. (2020). 'Joint statement on safety of journalists and access to information during the COVID-19 crisis'. 15 April 2020. https://www.gov.uk/government/news/joint-statement-on-safety-of-journalists-and-access-to-information-during-the-covid-19-crisis--2

UK Department for International Development. (2020). 'UK aid to tackle global spread of coronavirus 'fake news''. 12 March 2020. https://www.gov.uk/government/news/uk-aid-to-tackle-global-spread-of-coronavirus-fake-news

UK Government. (2019). 'Conflict, Stability and Security Fund: programme summaries for Eastern Europe, Central Asia and Western Balkans 2019 to 2020'. 05 November 2019. https://www.gov.uk/government/publications/conflict-stability-and-security-fund-programme-summaries-for-eastern-europe-central-asia-and-western-balkans-2019-to-2020

UK Government. (2020). 'UK aid to tackle global spread of coronavirus 'fake news'. Press Release, 12 March 2020. https://www.gov.uk/government/news/uk-aid-to-tackle-global-spread-of-coronavirus-fake-news

UK House of Commons Foreign Affairs Committee. (2019). '"Media freedom is under attack": The FCO's defence of an endangered liberty'. 04 September 2019. https://publications.parliament.uk/pa/cm201719/cmselect/cmfaff/1920/1920.pdf

UK ICO. (2018). *ICO Issues Maximum £500,000 Fine to Facebook for Failing to Protect Users' Personal Information*. Press Release, 25 October 2018. https://ico.org.uk/facebook-fine-20181025

UK MOD (Ministry of Defence), Defence Science and Technology Laboratory, Guto Bebb, and The Rt Hon Gavin Williamson CBE MP. (2018). 'Flagship AI Lab announced as Defence Secretary hosts first meet between British and American defence innovators'. Press Release, 22 May 2018. https://www.gov.uk/government/news/flagship-ai-lab-announced-as-defence-secretary-hosts-first-meet-between-british-and-american-defence-innovators

UK Parliament, Subcommittee on Online Harms and Disinformation. (2020). 'Disinformation and misinformation on social media about COVID-19 to be investigated'. 26 March 2020. https://committees.parliament.uk/committee/438/subcommittee-on-online-harms-and-disinformation/news/145761/disinformation-and-misinformation-on-social-media-about-covid19-to-be-investigated/

UN. (1948). 'Universal Declaration of Human Rights'. General Assembly Resolution 217A, 10 December 1948. https://www.un.org/en/universal-declaration-human-rights/

UN. (2019). United Nations Strategy and Plan of Action on Hate Speech. May 2019. https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml

UN Africa Renewal. (2020). 'Mauritius, Senegal, South Africa among authors of global call against COVID-19 'infodemic'. 22 June 2020. https://www.un.org/africarenewal/news/coronavirus/cross-regional-statement-%E2%80%9Cinfodemic%E2%80%9D-context-covid-19

UN Department of Global Communications. (2020). '5 ways the UN is fighting 'infodemic' of misinformation'. 30 April 2020. https://www.un.org/en/un-coronavirus-communications-team/five-ways-united-nations-fighting-%E2%80%98infodemic%E2%80%99-misinformation

UN Human Rights. (2018). 'UN experts call on India to protect journalist Rana Ayyub from online hate campaign'. https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=23126&LangID=E

UN Human Rights. (2020a). 'COVID-19: Governments must promote and protect access to and free flow of information during pandemic – International experts. https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25729&LangID=E

UN Human Rights. (2020b). 'Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development'. 14 July 2020. A/HRC/44/L.18/Rev.1 https://undocs.org/A/HRC/44/L.18/Rev.1

UN News. (2020). 'Hatred going viral in 'dangerous epidemic of misinformation' during COVID-19 pandemic'. 14 April 2020. https://news.un.org/en/story/2020/04/1061682

UN Secretary General. (2020). 'Secretary-General's remarks to High-Level Dialogue on Press Freedom and Tackling Disinformation in the COVID-19 Context [bilingual, as delivered; scroll down for English and French versions]'. 04 May 2020. https://www.un.org/sg/en/content/sg/statement/2020-05-04/secretary-generals-remarks-high-level-dialogue-press-freedom-and-tackling-disinformation-the-covid-19-context-bilingual-delivered-scroll-down-fo-r-english-and-french

UN Special Rapporteur on Freedom of Opinion and Expression et. al. (2017). 'Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda'. UN Document FOM.GAL/3/17, https://www.osce.org/fom/302796?download=true

UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression. (2018a). Report of the Special Rapporteur to the General Assembly on AI and its impact on freedom of opinion and expression. A/73/348. https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportGA73.aspx

UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression. (2018b). Report to the United Nations Human Rights Council on A Human Rights Approach to Platform Content Regulation, A/HRC/38/35, https://freedex.org/wp-content/blogs.dir/2015/files/2018/05/G1809672.pdf

UNESCO. (2016). 'UNESCO launches Countering Online Hate Speech publication'. https://unesdoc.unesco.org/ark:/48223/pf0000233231

UNESCO. (2019). 'Elections and media in digital times'. https://en.unesco.org/sites/default/files/in_focus_world_trends_report_2019_brochure_final_2.pdf

UNESCO. (2020). 'Journalism, press freedom and COVID-19'. https://en.unesco.org/sites/default/files/unesco_covid_brief_en.pdf

UNI Global Union. (2020). 'Solidarity: Fighting the COVID-19 Crisis'. https://www.uniglobalunion.org/sites/default/files/files/news/20_03_25_en_uni_mei_report_on_response_to_covid-19_in_the_media_enteratinment_sector.pdf

United Nations. (1994). 'Professional Training Series No. 2. Human Rights and Elections. A Handbook on the Legal, Technical and Human Rights Aspects of Elections. https://eos.cartercenter.org/uploads/document_file/path/8/training2enTCCoptimized.pdf

US Senate Select Committee on Intelligence. (2018). 'New Reports Shed Light on Internet Research Agency's Social Media Tactics'. Press Release, 17 December 2017. https://www.intelligence.senate.gov/press/new-reports-shed-light-internet-research-agency%E2%80%99s-social-media-tactics

US Senate Select Committee on Intelligence. (2019a). 'Russian Active Measures, Campaigns and Interference In the 2016 U.S. Election'. Volume 1, July 2019. https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume1.pdf

US Senate Select Committee on Intelligence. (2019b). 'Russian Active Measures, Campaigns and Interference In the 2016 U.S. Election'. Volume 2, October 2019. https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf

US Senate Select Committee on Intelligence. (2019c). 'Russian Active Measures, Campaigns and Interference In the 2016 U.S. Election'. Volume 3, February 2020. https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume3.pdf

US Senate Select Committee on Intelligence. (2019d). 'Senate Intel Committee Releases Bipartisan Report on Russia's Use of Social Media'. Press Release, 08 October 2019. https://www.intelligence.senate.gov/press/senate-intel-committee-releases-bipartisan-report-russia's-use-social-media

Vaccari, C. & Chadwick, A. (2020). 'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News'. *Social Media + Society*, 19 February 2020. https://doi.org/10.1177/2056305120903408

Van Den Berg, E. & Snelderwaard, T. (2019). 'Nu.nl stopt met bestrijding nepnieuws voor Facebook na terugdraaien factchecks PVV en FvD'. *NPO3*, 26 November 2019. https://www.npo3.nl/brandpuntplus/nu-nl-stopt-met-bestrijding-nepnieuws-voor-facebook-we-willen-ook-de-pvv-kunnen-factchecken

Van Dieman, E. (2019). '70 complaints of 'digital disinformation' over past month - Electoral Commission'. *News 24*, 06 May 2019. https://www.news24.com/elections/news/70-complaints-of-digital-disinformation-over-past-month-electoral-commission-20190506

Van Dijck, J., Poell, T. & de Waal, M. (2018). 'The Platform Society: Public Values in a Connective World'. Oxford University Press, 2018.

Varol, O., Ferrara, E., Davis, C.A., Menczer, F. & Flammini, A. (2017). 'Online Human-Bot Interactions: Detection, Estimation, and Characterization'. In Intl. AAAI Conf. on Web and Social Media (ICWSM), 09 March 2017. https://arxiv.org/abs/1703.03107

Verberne, S. (2018). 'Explainable IR for Personalizing Professional Search'. In Joint Proceedings of the 1st International Workshop on Professional Search (ProfS2018), co-located with SIGIR. http://webcache.googleusercontent.com/search?q=cache:GeXr_yAHVGYJ:ceur-ws.org/Vol-2127/paper4-profs.pdf+&cd=1&hl=en&ct=clnk&gl=uk&client=firefox-b-e

Verdoliva, L. (2020). 'Media Forensics and DeepFakes: an overview'. https://arxiv.org/abs/2001.06564

Vicol, D-O. (2020). 'What makes us believe a false claim? Age, education, and cognitive biases all play a part'. *Fullfact*, 28 February 2020. https://fullfact.org/blog/2020/feb/what-makes-us-believe-false-claim/

Viner, K. (2017). 'A mission for journalism in a time of crisis'. *The Guardian*, 17 November 2017. https://www.theguardian.com/news/2017/nov/16/a-mission-for-journalism-in-a-time-of-crisis

Vishwanath, A. (2020). 'Explained: The laws being used to suspend Internet, and what SC laid down'. *The Indian Express*, 11 January 2020. https://indianexpress.com/article/explained/kashmir-supreme-court-internet-shutdown-sec-144-how-to-read-judgment-6209676/

Vlachos, A., & Riedel, S. (2015). 'Identification and Verification of Simple Claims about Statistical Properties'. In volume 'Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing', 17-21 September 2015, Lisbon, Portugal. Association for Computational Linguistics, 2596-2601. https://www.aclweb.org/anthology/D15-1312/

Vosoughi, S., Roy, D. & Aral, S. (2018). 'The spread of true and false news online'. *Science*: Vol. 359, Issue 6380, pp. 1146-1151

Waddell, K. (2019) 'A digital breadcrumb trail for deepfakes.' *AXIOS*, 12 July 2019. https://www.axios.com/deepfake-authentication-privacy-5fa05902-41eb-40a7-8850-5450bcad0475.html

Walker, S., Mercea, D. & Bastos, M. (2019). 'The disinformation landscape and the lockdown of social platforms'. https://doi.org/10.1080/1369118X.2019.1648536

Wang, A.H. (2010). 'Don't follow me: Spam detection in Twitter'. *In Proceedings of the International Conference on Security and Cryptography* (*SECRYPT 2010*). https://ieeexplore.ieee.org/abstract/document/5741690

WAN-IFRA. (2016). 'The 2016 Global Report on Online Commenting: Executive Summary'. 17 October 2016. https://blog.wan-ifra.org/2016/10/17/the-2016-global-report-on-online-commenting-executive-summary

WARC. (2018). '65% of digital media will be programmatic in 2019'. 19 November, 2018. https://www.warc.com/newsandopinion/news/65_of_digital_media_will_be_programmatic_in_2019/41341

Wardle, C. (2017a). 'Fake News. it's Complicated'. *First Draft*, 16 February 2017. https://firstdraftnews.org/fake-news-complicated/,

Wardle, C. (2017b). 'Foreword in The Impact of CrossCheck on Journalists & the Audience'. *First Draft*. https://firstdraftnews.org/wp-content/uploads/2017/11/Crosscheck_rapport_EN_1129.pdf/

Wardle, C. & Derakhshan, H. (2017). 'Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making'. Council of Europe report DGI, 09 October 2017. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Warner, M. R. (2017). 'The Honest Ads Act'. US Senator of Virginia, US, May 2019. https://www.warner.senate.gov/public/index.cfm/the-honest-ads-act

Wasserman, H. (2020). 'Laughter in the time of a pandemic: why South Africans are joking about coronavirus'. *The Conversation*, 15 March 2020. https://theconversation.com/laughter-in-the-time-of-a-pandemic-why-south-africans-are-joking-about-coronavirus-133528

Wasserman, H. & Madrid-Morales, D. (2018). 'Study sheds light on scourge of "fake" news in Africa'. The Conversation, 21 November 2018. https://theconversation.com/study-sheds-light-on-scourge-of-fake-news-in-africa-106946

WebWire. (2020). 'NEW REPORT: YouTube found promoting climate denial to millions'. 20 January 2020. https://www.webwire.com/ViewPressRel.asp?aId=253667

Weedon, J., Nuland, W. & Stamos, A. (2017). 'Information Operations and Facebook'. *Facebook*, 27 April 2017, https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf

Weinberger, D. (2009). 'Transparency is the new objectivity'. https://www.hyperorg.com/blogger/2009/07/19/transparency-is-the-new-objectivity/

WhatsApp. (2019a). 'More changes to forwarding'. *WhatsApp Blog*, 21 January 2019. https://blog.whatsapp.com/more-changes-to-forwarding

WhatsApp. (2019b). 'Stopping Abuse: How WhatsApp Fights Bulk Messaging and Automated Behavior'. https://scontent.whatsapp.net/v/t61/69510151_652112781951150_6923638360331596993_n.pdf/Stopping-Abuse-white-paper.pdf?_nc_sid=2fbf2a&_nc_ohc=hwbXZHmgl3oAX9FN2oR&_nc_ht=scontent.whatsapp.net&oh=c5105a645efd83ac2e78d5e3e33247c3&oe=5E61E6EE

WHO. (2018). 'Risk Communication and Community Engagement (RCCE) Considerations: Ebola Response in the Democratic Republic of the Congo'. May 2018. https://www.afro.who.int/publications/risk-communication-and-community-engagement-rcce-considerations-ebola-response

WHO. (2020). Coronavirus, 13 February 2020. Press Conference Transcript: https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-full-press-conference-13feb2020-final.pdf?sfvrsn=b5435aa2_2

Who Targets Me. (2019). 'Despite protestations from the @bbc, the @Conservatives haven't stopped running the ad.' *Twitter*, 30 November 2019. https://twitter.com/WhoTargetsMe/status/1200729905408987136

Wilding, D., Fray, P., Molitorisz, S. & McKewon, E. (2018). 'The Impact of Digital Platforms on News and Journalistic Content'. University of Technology Sydney. https://www.accc.gov.au/system/files/ACCC%20commissioned%20report%20-%20The%20impact%20of%20digital%20platforms%20on%20news%20and%20journalistic%20content%2C%20Centre%20for%20Media%20Transition%20%282%29.pdf

Wilner, T. (2018). 'We can probably measure media bias. But do we want to?' *Columbia Journalism Review*, 9 January 2018. https://www.cjr.org/innovations/measure-media-bias-partisan.php

Witness Media Lab. (2019). 'Ticks or it didn't happen: Confronting key dilemmas in authenticity infrastructure for multimedia'. December 2019. https://lab.witness.org/ticks-or-it-didnt-happen/

Wong, J. C. (2019a). 'Google latest tech giant to crack down on political ads as pressure on Facebook grows'. *The Guardian*, 21 November 2019. https://www.theguardian.com/technology/2019/nov/20/google-political-ad-policy-facebook-twitter

Wong, J. C. (2019b). 'Sri Lankans fear violence over Facebook fake news ahead of election'. *The Guardian*, 12 November 2019. https://www.theguardian.com/world/2019/nov/11/facebook-sri-lanka-election-fake-news

Wong, J. C. (2020a). 'Will Facebook's new oversight board be a radical shift or a reputational shield?'. *The Guardian*, 07 May 2020. https://www.theguardian.com/technology/2020/may/07/will-facebooks-new-oversight-board-be-a-radical-shift-or-a-reputational-shield

Wong, J. C. (2020b). 'Facebook removes Trump post over false Covid-19 claim for first time'. *The Guardian*, 06 August 2020. https://www.theguardian.com/technology/2020/aug/05/facebook-donald-trump-post-removed-covid-19

Wood, A. K. & Ravel, A.M. (2018). 'Fool me once: regulating "fake news" and other online advertising'. *Southern California Law Review*. Vol. 91. Pages 1223-1278. https://southerncalifornialawreview.com/wp-content/uploads/2018/10/91_6_1223.pdf

Woolley, S.C. & Howard, P.N. (2016). 'Automation, Algorithms, and Politics: Political Communication, Computational Propaganda, and Autonomous Agents. *International Journal of Communication*, 10 October 2016. http://ijoc.org/index.php/ijoc/article/view/6298/1809

World Summit Working Group. (2005). 'Report of the Working Group on Internet Governance'. United Nations & World Summit, June 2005. https://www.wgig.org/docs/WGIGREPORT.pdf

Wu, T. (2017). 'Blind Spot: The Attention Economy and the Law'. *Antitrust Law Journal, Forthcoming*, 26 March 2017. https://ssrn.com/abstract=2941094 or http://dx.doi.org/10.2139/ssrn.2941094

Wulczyn, E., Thain, N., & Dixon, L. (2017). 'Ex machina: Personal attacks seen at scale'. *In Proceedings of the 26th International Conference on World Wide Web, 1391–1399*, 25 February 2017. https://arxiv.org/abs/1610.08914

Xinhua. (2019). 'Consensus issued after fourth meeting of World Media Summit presidium in Shanghai'. 20 November 2019. http://www.xinhuanet.com/english/2019-11/20/c_138567780.htm

Yang, Z. (2018). '网信办公布新规 微博客服务提供者应建立健全辟谣机制'. *People's Daily*, 03 February 2018. http://www.xinhuanet.com/2018-02/03/c_1122362053.htm

Zacharia, J. (2019). 'Time is running out to to fight disinformation in 2020 election'. *San Francisco Chronicle*, 31 August 2019. https://www.sfchronicle.com/opinion/article/Time-is-running-out-to-to-fight-disinformation-in-14404399.php?psid=ms95Y

Zadrozny, B. (2019). 'Anti-vaccination groups still crowdfunding on Facebook despite crackdown'. *NBC News*, 11 October 2019. https://www.nbcnews.com/tech/internet/anti-vaccination-groups-still-crowdfunding-facebook-despite-crackdown-n1064981

Zampoglou, Z. Papadopoulos, S., Kompatsiaris, Y., Bouwmeester, R., & Spangenberg, J. (2016). 'Web and Social Media Image Forensics for News Professionals'. *In Proceedings of the Tenth International AAAI Conference on Web and Social Media*, 16 April 2016. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13206

Zarocostas, J. (2020). 'How to fight an infodemic'. The Lancet, 29 February 2020. https://doi.org/10.1016/S0140-6736(20)30461-X

Zeit. (2019). 'Deutsche Behörde verhängt Millionenstrafe gegen Facebook'. 02 July 2019. https://www.zeit.de/digital/datenschutz/2019-07/facebook-hasskommentare-fake-news-millionenstrafe-bussgeld-netzdg

Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). 'Defending Against Neural Fake News'. 29 May 2019. http://arxiv.org/abs/1905.12616

Zharov, A. (2019). 'Russia to Set Up 'Fake News Database''. *The Moscow Times*, 16 May 2019. https://www.themoscowtimes.com/2019/05/16/russia-to-set-up-fake-news-database-a65613

Zilavy, T. (2018). 'What is Hashing and How Does It Work?'. Medium, 06 October 2018. https://medium.com/datadriveninvestor/what-is-hashing-and-how-does-it-work-7800f461a0de

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G., Tolmie, P. (2016). 'Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads'. PLOS ONE 11(3): e0150989. https://doi.org/10.1371/journal.pone.0150989

Zuboff, S. (2019). 'The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power'. New York: Public Affairs

Zuckerberg, M. (2016a). 'Announcement from Zuckerberg regarding the launch of a "third party verification" program'. *Facebook*, 18 November 2016. https://www.facebook.com/zuck/posts/10103269806149061

Zuckerberg, M. (2016b). 'I want to share some thoughts about Facebook and the election'. https://www.facebook.com/zuck/posts/i-want-to-share-some-thoughts-on-facebook-and-the-electionour-goal-is-to-give-ev/10103253901916271/

Zuckerberg, M. (2018). Testimony before Senate: "Why do you shift the burden to users to flag inappropriate content and make sure it's taken down?". *BuzzFeed*, 10 April 2018. https://twitter.com/BuzzFeedNews/status/983815663469948929

5Rights Foundation. (2019). 'Towards an Internet Safety Strategy'. *5Rights Foundation*, January 2019. https://5rightsfoundation.com/uploads/final-5rights-foundation-towards-an-internet-safety-strategy-january-2019.pdf

10

المُـلحق (أ)



التحقيقات وفرق العمل والمبادئ التوجيهية

مع تحوّل المعلومات المغلوطة واسعة الانتشار إلى مصدر قلق متزايد، شكّلت بلدان عديدة فرق عمل مكرّسة لهذا الموضوع، عملت على رصد حملات التّضليل والتحقيق فيها. وكثيراً ما انطلق عمل فرق العمل هذه في أعقاب حملات تضليل اعتُبرت تهديداً للسلامة الديمقراطية أو للأمن السيبراني في البلد. وتمحور النطاق الأوّلي لهذه المبادرات الحكومية حول البُعد التثقيفي، حيث شملت 24 مبادرة من أصل 33 جانباً يتعلّق بالدراية الإعلامية. بالإضافة إلى ذلك، تضمّنت 17 مبادرة في هذه الفئة التحقُّق من الوقائع. ولابدّ من الإشارة في هذا الصدد إلى أنه من أصل 33 بلداً باشرت بهكذا تحقيقات أو أنشأت فرق عمل من هذا القبيل، ركّز 21 منها على الانتخابات. وتهدف التحقيقات المركّزة على الانتخابات إلى التحقيق في التدخلات في العملية التشريعية أو إلى منعها. ونظراً إلى أن المعلومات المُضلّلة الإلكترونية لظاهرة جديدة نسبياً، فإن معظم المبادرات التي تمّ التعرّف إليها حديثة العهد ولا تزال عُرضة للتطوُّر، أو قد تؤدي في نهاية المطاف إلى مبادرات تنظيمية.

1. إطار عمل رابطة أمم جنوب شرق آسيا (آسيان) وإعلانها المشترك للحد من الآثار الضارة للأخبار الزائفة (2018)

أصدر الوزراء المسؤولون عن الإعلام (ASEAN AMRI, 2018) في رابطة أمم جنوب شرق آسيا «إعلاناً مشتركاً بشأن إطار العمل الرامي إلى تقليص الآثار الضارة للأخبار الزائفة». وقد شجّعوا فيه على تبادل أفضل الممارسات بين أعضاء الرابطة، واقترحوا إطار عمل يستند إلى أربعة أُسُس هي: التعليم والتوعية (الإلمام الرقمي والمسؤولية الخاصة)، والكشف والاستجابة (رصد الحكومة وكشف المعلومات المُضلّلة، ورصد وسائل الإعلام والتحقُّق من الوقائع)؛ والمجتمع المحلي والمشاركة من أسفل إلى أعلى (كشف المواطن والمجتمع المدني للتضليل)؛ والقواعد والمبادئ التوجيهية (حسب السياق والقوانين والقواعد والمبادئ التوجيهية الوطنية من أجل «تمكين المواطنين وحمايتهم» و«الترويج لاستحداث وتبادل مسؤولَين للمعلومات على الإنترنت») (ASEAN AMRI, 2018).

2. فريق العمل الأُسترالي لضمان النزاهة الانتخابية (2019)

أنشأت اللجنة الانتخابية الأسترالية (AEC, 2019a) فريق عمل مع الوكالات الحكومية الأخرى في الفترة التي سبقت الانتخابات الاتحادية في أيار/مايو 2019، من أجل حماية نزاهة الانتخابات، من خلال ابتكار حلول تركّز على الإذن بالتواصل الانتخابي، والشفافية لفائدة التأثير الأجنبي، والأمن السيبراني. ونظّمت اللجنة أيضاً حملة للدراية الإعلامية عنوانها «تأنَّ وفكّر» (AEC, 2019b). على ذلك، أبرزت عدة تقارير حكومية تتعلق بوسائل الإعلام وتتناول على وجه الخصوص مستقبل الصحافة ذات المصلحة العامة وتأثير المنصات الرقمية على المنافسة الإعلامية، والحاجة إلى الشفافية والدراية الإعلامية في مُكافحة المعلومات المُضلّلة (Buckmaster & Wils, 2019).

3. اللجنة الدائمة المشتركة المعنية بالمسائل الانتخابية في البرلمان الأسترالي: الديمقراطية والتّضليل (2018)

أشارت اللجنة الدائمة المشتركة المعنية بالمسائل الانتخابية في البرلمان الأسترالي (2019) إلى وجود معلومات مضللة في الانتخابات الأمريكية لعام 2016، وفي الاستفتاء حول خروج بريطانيا من الاتحاد الأوروبي في المملكة المتحدة، علاوةً على توفّر أدلة على المعلومات المُضلّلة في العمليات الديمقراطية على الصعيد الدولي في عام 2018، وقررت إدراج «الديمقراطية والمعلومات المُضلّلة ضمن الرقابة المستمرة التي تمارسها على المسائل الانتخابية الأسترالية، مع التركيز على المجالات التالية:

- مدى استهداف روبوتات التواصل الاجتماعي للناخبين الأستراليين والخطاب السياسي في الماضي؛
- المصادر المحتملة للتلاعب بوسائل التواصل الاجتماعي داخل أستراليا وعلى الصعيد الدولي؛
- طرق معالجة انتشار «الأخبار الزائفة» عبر الإنترنت خلال الحملات الانتخابية؛
- التدابير الآيلة إلى تحسين الدراية الإعلامية للناخبين الأستراليين».

4. فريق الخبراء والمنصة التشاركية في بلجيكا (2018)

أنشأ وزير الأجندة الرقمية البلجيكي فريق خبراء (Alaphilippe et.al.,2018a)، ومنصة تشاركية (Monopinion.belgium.be) [488]، وحلقة نقاش للمواطنين لمناقشة موضوع المعلومات المُضلّلة في عام 2018. ونتيجةً لهذه المشاورات، قدّمت الحكومة الاتحادية مبلغ 1.5 مليون يورو لدعم مبادرات الدراية الإعلامية التي تزيد من شفافية النظم الإيكولوجية الرقمية (مثل شفافية الإعلانات) والقدرة على إيجاد معلومات عالية الجودة ومتنوعة (مثل مؤشرات المصدر) (De Croo, 2018).

5. المحكمة الانتخابية العليا في البرازيل (2018)

وقّعت المحكمة الانتخابية العليا في البرازيل مذكرة تفاهم مع غوغل وفيس بوك للحدّ من انتشار المعلومات المُضلّلة الانتخابية في الفترة السابقة للانتخابات العامة في تشرين الأول/أكتوبر 2018 (Alves, 2018).

6. تقرير اللجنة البرلمانية الكندية حول الديمقراطية المُهدَّدة (2018)

في أعقاب فضيحة Cambridge Analytica، نشرت اللجنة الدائمة المعنية بالوصول إلى المعلومات والخصوصية والأخلاقيات في مجلس العموم الكندي (2018) دراسة حول المخاطر والحلول المتعلقة بالديمقراطية في عصر المعلومات المُضلّلة واحتكار البيانات، شملت مجموعة واسعة من التوصيات تطرقت إلى تطبيق التشريعات المتعلقة بالخصوصية، والأحزاب السياسية، وزيادة صلاحيات مفوض الخصوصية، وتنظيم منصات وسائل التواصل الاجتماعي والأمن السيبراني، فضلاً عن دعم الإلمام الرقمي والبحث في تأثير المعلومات المُضلّلة.

[488] https://monopinion.belgium.be/processes/stopfakenews

7. مبادرة الحكومة الكندية للمواطنين الرقميين (2019)

أنشأت الحكومة الكندية مُبادرة «المواطن الرقمي» قبل الانتخابات الفدرالية في تشرين الأول/أكتوبر 2019، «لبناء قدرة المواطنين على مواجهة المعلومات المُضلّلة عبر الإنترنت، وبناء شراكات لدعم نظام بيئي صحي للمعلومات» (Canadian Government, 2019b Canadian Government, 2019a;). وأسفرت هذه المبادرة عن استثمار كبير في الاستجابات لأزمة المعلومات المُضلّلة (وقد تراوحت بين البحث والتثقيف الصحفي، والتدريب على الدراية الإعلامية).

8. البروتوكول العام للحكومة الكندية حول الحوادث الانتخابية الخَطِرة (2019)

أرسى البروتوكول العام للحوادث الانتخابية الخطرة[489] «عملية بسيطة وواضحة ومحايدة» لإخطار الكنديين بأي تهديد يطال نزاهة الانتخابات العامة التي نُظّمت في العام 2019. وتألفت المجموعة المسؤولة عن البروتوكول (وهي تغطي الأمن الوطني والشؤون الخارجية والحكم الديمقراطي والمنظور القانوني) من كبار الموظفين العموميين الكنديين ذوي الخبرة، وأُنيطت بهم مسؤولية تحديد ما إذا كانت عتبة إبلاغ الكنديين تأمّنت.

9. تقرير مجلس أوروبا عن اضطراب المعلومات (2017)

في تشرين الأول/أكتوبر 2017، نشر مجلس أوروبا تقريراً بعنوان «اضطراب المعلومات: نحو إطار متعدد التخصصات للأبحاث وصُنع السياسات» (Wardle & Derakhshan, 2017). يميّز التقرير هذا بين المعلومات المغلوطة («عندما يتم تبادل معلومات كاذبة، ولكن لا يُقصد بها إلحاق الضرر»)، والمعلومات المُضلّلة («عندما يتم مشاركة معلومات كاذبة عن علم للتسبّب في الضرر») و«المعلومات الضارّة» («عندما يتم تبادل معلومات حقيقية للتسبّب في ضرر، وغالباً ما يتم ذلك عن طريق نشر معلومات مصمّمة لتبقى خاصة علناً»). ويقدم هذا التقرير توصيات لشركات التكنولوجيا والحكومات الوطنية ووسائل الإعلام والمجتمع المدني ووزارات التعليم. كما يؤمّن مجلس أوروبا موارد للدراية الإعلامية والمعلوماتية للتصدي للمعلومات المغلوطة والمُضلّلة والدعاية في إطار برنامجه التثقيفي المعنون «حُريّة الكلام.. سلامة التعلّم» Free to speak, Safe to learn[490].

10. خطة عمل الدنمارك للانتخابات (2018)

في عام 2018، كشفت وزارة الخارجية الدنماركية عن خطة عمل لمُكافحة التدخل الأجنبي المعادي في الانتخابات الدنماركية. وتتراوح المبادرات الإحدى عشرة بين إنشاء فريق عمل حكومي دولي للتنسيق، ورفع حالة التأهب في مجال الأمن السيبراني، وتدريب الأحزاب السياسية، والحوار مع وسائل الإعلام وشركات التواصل الاجتماعي بشأن نماذج التعاون الممكنة لمواجهة المحاولات المحتملة لممارسة التأثير الأجنبي. كما ستقوم الحكومة الدنماركية بتحديث القانون الجنائي لحماية الدنمارك من «التهديد الذي تمثّله حملات النفوذ التي تشنّها أجهزة الاستخبارات الأجنبية».

---

489 https://www.canada.ca/en/democratic-institutions/services/protecting-democracy/critical-election-incident-public-protocol.html

490 https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propagandamisinformation-and-fake-news

11. رابطة الدفاع السيبراني في إستونيا (2007)

لطالما استُهدفت إستونيا بالدعاية والمعلومات المُضلّلة عبر التاريخ. وبعد هجوم إلكتروني مكثّف في عام 2007، أنشأت إستونيا وحدة سيبرانية تابعة لرابطة الدفاع تقضي مهمّتها بـ«حماية أسلوب حياة إستونيا عالي التقنية، بما في ذلك حماية البنى التحتية للمعلومات ودعم أهداف الدفاع الوطني الأوسع نطاقاً»[491]. وتضمّ الوحدة مواطنين متطوعين مستعدين للمساهمة في استراتيجية الأمن السيبراني في البلاد. ومنذ عام 2008، تضمّ إستونيا مركز التميّز التعاوني للدفاع السيبراني التابع لحلف شمال الأطلسي[492] (Thompson, 2019).

12. فريق الخبراء رفيع المستوى ومدونة الممارسات، وخطة العمل بشأن المعلومات المُضلّلة التابعة للاتحاد الأوروبي (2018)

في آذار/مارس 2018، أصدر فريق خبراء رفيع المستوى يشمل أصحاب مصلحة متعددين بشأن الأخبار الزائفة والتّضليل عبر الإنترنت (EU HLEG, 2018) توصيات حول المعلومات المُضلّلة، مُركّزاً بشكل أساسي على الدور الذي يمكن أن تؤديه منصات التواصل الاجتماعي في دعم النظام البيئي الإعلامي، والتحقُّق من الوقائع، وتعزيز الدراية.

وقد تضمّن بلاغ متابعة صدر عن المفوضية الأوروبية (European Commission, 2018a) توصيات بشأن الدراية الإعلامية والتعددية، وإن جاء بشكل مقتضب. وأضافت المفوضية أيضاً أفكاراً بشأن العمليات الانتخابية والتواصل الاستراتيجي. كما أن المفوضية أخذت بأفكار فريق الخبراء رفيع المستوى حول النظام البيئي الرقمي الشفاف، فأنشأت منتدى الاتحاد الأوروبي لأصحاب المصلحة المتعددين من أجل تطوير مدونة ممارسات خاصة بالاتحاد الأوروبي (European Commission, 2018b).

تركّز مدونة الممارسات هذه (European Commission, 2018c) على الإعلانات (الانتخابية)، وتتضمّن قسماً موجزاً عن روبوتات الويب المميكنة، وتتناول دور المنصات في دعم/تمكين الدراية والتحقُّق من الوقائع والبحث. وتجرد مدونة الممارسات التدابير القائمة، من دون أن تسعى إلى توفير معايير خاصة بهذا القطاع. وبعد عام على وضع مدونة الممارسات، يخضع تنفيذها لاستعراض مستقل (European Commission, 2019).

تندرج كل هذه النقاط ضمن خطة العمل الأوسع للاتحاد الأوروبي بشأن التّضليل التي تهدف إلى بناء القدرات والتعاون داخل الاتحاد الأوروبي وبين دُوَلِه الأعضاء (European Commission and High Representative, 2018). ويدير القسم الأوروبي للعمل الخارجي أيضاً موقعاً على شبكة الإنترنت يهدف إلى تقديم سرديات مضادة للتضليل.

....

491 https://www.kaitseliit.ee/en/cyber-unit

492 https://ccdcoe.org/

13. مدونة الأخلاقيات الخاصة بمنصات التواصل الاجتماعي والمتعلّقة بالانتخابات العامة لعام 2019 في الهند (2019)

بناءً على تقرير اللجنة المعنية بالقسم 126 من قانون تمثيل الشعب لعام 1951 (تقرير لجنة سينها–Sinha) الذي تضمّن توصيات بشأن الحفاظ على الصمت الانتخابي خلال الساعات الثماني والأربعين التي تسبق الاقتراع (ECI, 2019)، وافقت منصات التواصل الاجتماعي ورابطة الإنترنت والهاتف المحمول في الهند (IAMAI, 2019) على مدونة أخلاقيات طوعية للانتخابات العامة لعام 2019. وتعهّدت العديد من شركات التواصل الاجتماعي بتنفيذ حملات توعية للناخبين، من أجل «إنشاء آلية إبلاغ ذات أولوية عالية» للجنة الانتخابية في الهند. واتفقت أيضاً على تعيين فرق مخصّصة للانتخابات، بناءً على إخطار من اللجنة الانتخابية، تزيل المُحتوى المُبلغ عنه في غضون ثلاث ساعات خلال فترة الصمت الانتخابي التي تدوم 48 ساعة قبل الاقتراع، لتوفير آلية للمعلنين السياسيين لتقديم شهادات مسبقة تصدر عن اللجنة الانتخابية واتخاذ إجراءات بشأن الطلبات هذه على وجه السرعة، وتسهيل الشفافية في الإعلانات السياسية مدفوعة الأجر (IAMAI, 2019).

14. «غرفة العمليات» وحملة «وقف الخِدع» في إندونيسيا (2019)

في الفترة التي سبقت الانتخابات الرئاسية في نيسان/أبريل 2019، أنشأت وزارة الاتصالات وتكنولوجيا المعلومات الإندونيسية «غرفة عمليات» للكشف عن المُحتوى السلبي الذي ينتهك القواعد، وتعطيله (Board, 2019).

وتعاونت الوزارة أيضاً مع الجمعية الإندونيسية لمُكافحة التشهير، ومبادرة أخبار غوغل، والكثير من منظمات المجتمع المدني لإدارة حملة لتعزيز الدراية عنوانها «وقف الخِدع»، استهدفت في المقام الأول الطلاب والنساء العاملات في المنزل. وتهدف ورش العمل، التي استمرت من آب/أغسطس إلى تشرين الأول/أكتوبر 2019، إلى مساعدة المشاركين على الكشف عن المعلومات المُضلّلة والخِدع.

15. اللجنة الدولية الكبرى حول التّضليل و«الأخبار الزائفة» (2018)

رأت اللجنة الدولية الكبرى المعنية بالتّضليل و«الأخبار الزائفة» النور في أعقاب تحقيق برلماني بريطاني في المعلومات المُضلّلة والأخبار الزائفة. تضمّ هذه اللجنة برلمانيين، وقد استمعت حتى الآن إلى أدلة تتعلق بالتّضليل، والبيانات الضخمة، وتنظيم المُحتوى الضار، والتدخل الانتخابي عبر الإنترنت. وعُقدت ثلاثة اجتماعات في لندن (تشرين الثاني/نوفمبر 2018، استضافتها اللجنة البرلمانية للمسائل الرقمية والثقافة والإعلام والرياضة في المملكة المتحدة)، وأوتاوا (أيار/مايو 2019، استضافتها اللجنة البرلمانية الدائمة الكندية؛ المعنية في المقام الأول بالوصول إلى المعلومات والخصوصية والأخلاقيات) ودبلن (تشرين الثاني/نوفمبر 2019، استضافتها اللجنة الآيرلندية المشتركة بشأن التواصل والعمل المناخي والبيئة).

16. المجموعة الآيرلندية المشتركة بين الإدارات المعنية بأمن العملية الانتخابية والتّضليل (2017)

في كانون الأول/ديسمبر 2017، أنشأت الحكومة الآيرلندية مجموعة مشتركة بين الإدارات لتقييم التهديدات التي تحيط بالعملية الانتخابية الآيرلندية. وفي التقرير الأول للمجموعة، الذي نُشر في تموز/يوليو 2018، كان الاستنتاج الرئيسي أنّ «المخاطر التي تُهدّد العملية الانتخابية في آيرلندا منخفضة نسبياً، ولكن انتشار المعلومات المُضلّلة عبر الإنترنت وخطر الهجمات الإلكترونية على النظام الانتخابي باتا يُشكّلان مخاطر أكبر» (Irish IDG, 2019). وأوجز التقرير سبع توصيات، بما في ذلك إنشاء لجنة انتخابية، ومواصلة مبادرات الدراية الإعلامية، وتعزيز تدابير الأمن السيبراني. وفي تشرين الثاني/نوفمبر 2019، قدمت الحكومة الآيرلندية اقتراحاً تشريعياً لتنظيم شفافية الإعلانات السياسية المدفوعة عبر الإنترنت خلال فترات الانتخابات، تماشياً مع توصية أخرى صدرت عن المجموعة المشتركة بين الإدارات (Irish Department of the Taoiseach, 2019).

17. حملة «كفى خدعاً hoaxes» (2017) وبوابة «الزر الأحمر» في إيطاليا (2018)

في تشرين الأول/أكتوبر 2017، أعلنت وزارة التربية الإيطالية (2017) عن إطلاق حملة عنونتها «كفى خدعاً» (Bastabufale.it)[493] لتعزيز الدراية الإعلامية في المدارس الابتدائية والثانوية. وقبل الانتخابات العامة في آذار/مارس 2018، تم إطلاق بوابة «الزر الأحمر»[494] التي تتيح للمواطنين إبلاغ وحدة سيبرانية خاصة تابعة للشرطة عن المعلومات المُضلّلة. وستقوم وحدة الشرطة بالتحقيق في المُحتوى، ومساعدة المواطنين على الإبلاغ عن المعلومات المُضلّلة على منصات التواصل الاجتماعي، ورفع دعوى قضائية في حال كان المُحتوى تشهيرياً أو غير قانوني (la Cour, 2019). إلى جانب ذلك، نشرت الهيئة الناظمة للاتصالات في إيطاليا AGCOM (Italian AGCOM, 2018) مبادئ توجيهية قبل الانتخابات لضمان المساواة في معاملة جميع الأحزاب السياسية وشفافية الإعلانات السياسية والتشجيع على التحقُّق من الوقائع عبر الإنترنت.

18. مجموعة الدراسة حول خدمات المنصات في اليابان (2018)

أنشأت وزارة الشؤون الداخلية والاتصالات اليابانية (2018) مجموعة لدراسة استخدام منصات التواصل الاجتماعي للبيانات الشخصية في تشرين الأول/أكتوبر 2018. وتندرج استجابات المنصات للمعلومات المُضلّلة في نطاق الدراسة أيضاً، مع الإشارة إلى مدونة الممارسات الخاصة بالاتحاد الأوروبي. وفي كانون الأول/ديسمبر 2019، أفادت «وكالة الأنباء اليابانية جيجي– Jiji» أنه سيُطلب من منصات التواصل الاجتماعي «الكشف عن معاييرها لإزالة الأخبار الزائفة من خلال الذكاء الاصطناعي أو تكنولوجيات أخرى وإنشاء أنظمة للتعامل مع الشكاوى». وكان من المتوقع أن يصدر التقرير النهائي لمجموعة الدراسة في أوائل عام 2020 (Jiji, 2019).

....

493 https://www.bastabufale.it

494 https://www.interno.gov.it/it/notizie/progetto-red-button-contro-fake-news

19. المعهد الانتخابي الوطني في المكسيك (2018-2019)

وقّع المعهد الانتخابي الوطني المكسيكي اتفاقية تعاون مع فيس بوك وتويتر وغوغل للحدّ من انتشار المعلومات المُضلّلة الانتخابية ولنشر معلومات عمليّة عن الانتخابات خلال انتخابات العامَين 2018 و2019. وتوفّر شركات التواصل الاجتماعي أيضاً التدريب للصحفيين والمسؤولين من المكتب الانتخابي. وخلال الانتخابات، قام المعهد الانتخابي الوطني أيضاً بالتحقُّق من الوقائع في المعلومات (Mexico INE, 2019). وأكّد الرئيس «كوردوفا فيانيللو» أنه «على ضوء المعضلة المتمثّلة في تنظيم استخدام المنصات والشبكات الاجتماعية أو الحدّ منه في الحملات السياسية، فضّل المعهد الانتخابي الوطني اختيار نموذج غير عقابي من خلال توفير معلومات صحيحة وحقيقية رداً على المعلومات المغلوطة، مع بذل جهد لتثقيف الجمهور وتعزيز قدرات المواطنين على التمييز بين نوعيات المعلومات ومصادرها» (ترجمة أجراها المؤلفون، Mexico INE, 2019).

20. حملة «تحلَّ بالحسّ الناقد» في هولندا (2019)

أطلقت وزارة الداخلية الهولندية حملة للدراية الإعلامية عنوانها «تحلَّ بالحسّ الناقد» (Dutch Government, 2019a) في الفترة التي سبقت انتخابات البرلمان الأوروبي في أيار/مايو 2019. وفي تشرين الأول/أكتوبر 2019، اعتمدت الحكومة الهولندية استراتيجية لمُكافحة التّضليل، من خلال التشديد على الدراية الإعلامية، وشفافية منصات التواصل الاجتماعي والأحزاب السياسية (التي يُفضّل أن تؤمَّن من خلال التنظيم الذاتي)، والحفاظ على المشهد متعدد الأشكال. وقد أعربت الحكومة الهولندية أيضاً في الاستراتيجية عن تأييدها الصريح لمدونة الممارسات التي وضعها الاتحاد الأوروبي بشأن التّضليل وفريق عمل EEAS East StratCom (Dutch Government, 2019b). ويُعتبر التحقُّق من الوقائع وسيلة مهمّة لمُكافحة المعلومات المُضلّلة، غير أنّ «التصدي لمُحتوى المعلومات المُضلّلة على هذا النحو ليس في المقام الأول، بحسب الحكومة، مهمّة للحكومات أو لمؤسسات الاتحاد الأوروبي بل للصحافة والعلم بشكل أساسي، سواء جرى ذلك بالتعاون مع خدمات الإنترنت أم من دونها» (ترجمة أجراها المؤلفون، Dutch Government, 2019c).

21. التحقيق البرلماني في الانتخابات العامة لعام 2017 والانتخابات المحلية لعام 2016 في نيوزيلندا (2019)

في كانون الأول/ديسمبر 2019، نشرت لجنة العدل البرلمانية في نيوزيلندا تقريرها عن التحقيق في انتخابات العامَين 2016 و2017. وفي إطار التحقيق واسع النطاق، أثارت القلق بشأن التدخل الأجنبي خلال الانتخابات ورفعت توصيات لمنع التأثير الأجنبي عبر الإعلان والقرصنة والدعاية الشعبية الزائفة والتّضليل ولتقليصه.

وأوصى التقرير أيضاً الحكومة بالنظر في إمكانية تطبيق التوصيات المقدمة بشأن المعلومات المُضلّلة في تقرير لجنة التحقيق في الشؤون الرقمية والثقافة والإعلام والرياضة في مجلس العموم البريطاني حول «المعلومات المُضلّلة» و«الأخبار الزائفة» (وهي توصيات تتضمّن إنشاء هيئة ناظمة مستقلة لرصد سلوك المنصات، وتحسين الدراية الإعلامية، وتحسين الشفافية في الإعلانات عبر الإنترنت)، وتقرير اللجنة الدائمة المشتركة الأسترالية المعنية بالمسائل الانتخابية حول «إجراء الانتخابات الاتحادية والمسائل المتصلة بها لعام 2016» (الذي يتضمّن إنشاء فريق عمل دائم لمنع التلاعب السيبراني ومكافحته، وزيادة الوضوح في الإطار التنظيمي على المنصات) (New Zealand Parliament Justice Committee, 2019).

وقبل أسبوع واحد من نشر التقرير النهائي للتحقيق، تم تقديم قانون وإقراره في البرلمان لحظر التبرعات السياسية الأجنبية خلال الانتخابات (انظر: «التشريعات المعتمدة») التي يمكن أن تؤثّر على مشهد التّضليل.

22. دليل منظّمة الدول الأمريكية بشأن حُريّة التّعبير والتّضليل المتعمّد في السياقات الانتخابية (2019)

قام المقرر الخاص المعني بحُريّة التّعبير التابع للجنة البلدان الأمريكية لحقوق الإنسان، وإدارة القانون الدولي وإدارة التعاون والمراقبة الانتخابيين التابعتان لمنظمة الدول الأمريكية بوضع دليل عمليّ «لضمان حُريّة التّعبير والنفاذ إلى المعلومات من مصادر مختلفة على شبكة الإنترنت في السياقات الانتخابية، دون تدخّل غير مبرّر» (OAS, 2019).

هذا الدليل يؤمّن توصيات لمجموعة واسعة من الجهات الفاعلة في الولايات الأمريكية كالسلطة التشريعية، والسلطة القضائية، والسلطة التنفيذية، والسُّلطات الانتخابية، ووسائط الإنترنت، والأحزاب السياسية، وشركات الاتصالات السلكية واللاسلكية، ووسائل الإعلام والصحفيين، والجهات التي تتحقّق من الوقائع، والشركات التي تتاجر بالبيانات لأغراض الدعاية، والجامعات، ومراكز الأبحاث (OAS, 2019). تتطرّق الصفحات المخصصة للتوصيات والتي يبلغ عددها 10 صفحات إلى تجنّب التشريعات التي تُحمّل وسائط الإنترنت مسؤولية المُحتوى الذي تنتجه أطراف ثالثة، وتعزيز التشريعات التي تؤمّن الحماية للمواطنين (حماية البيانات، وحُريّة التّعبير، وشفافية الإعلان الانتخابي)، وتعزيز النفاذ الشامل إلى الإنترنت، وحملات الدراية الإعلامية، والحماية الصحفية، والتعاون، والشفافية، ومراعاة الأصول القانونية عند مراقبة المُحتوى على الإنترنت (OAS, 2019).

23. مخزن إعلانات الأحزاب السياسية الرقمي وآلية الشكاوى المتعلقة بالتّضليل الرقمي (Real411) في جنوب افريقيا (2019)

أطلقت اللجنة الانتخابية في جنوب أفريقيا مبادرتَين في الفترة السابقة للانتخابات الوطنية والإقليمية في البلاد في عام 2019 من أجل معالجة موضوع التّضليل الرقمي وضمان وصول المواطنين إلى المعلومات خلال فترة الانتخابات. وشجّعت اللجنة الانتخابية أولاً الأحزاب السياسية على نشر إعلاناتها الرسمية في مخزن إعلانات الأحزاب السياسية الرقمي (padre.org.za) [495]. وأطلقت اللجنة ثانياً بوابة شكاوى للمواطنين للإبلاغ عن المعلومات المُضلّلة الرقمية خلال الانتخابات (real411.org) [496]. وقد نقلت اللجنة الانتخابية أنها تلقّت أكثر من 70 شكوى بشأن المعلومات المُضلّلة الرقمية في غضون شهر واحد من إطلاق المنصة (van Diemen, 2019). ومنذ ذلك الحين، وسّعت البوابة نطاق اختصاصها لتشمل أيضاً تلقّي الشكاوى المتعلقة بالجرائم الرقمية المتصلة بالتحريض على العنف وخطاب الكراهية ومُضايقة الصحفيين أيضاً. وتتلقى كلتا المبادرتين الدعم من مجموعة المجتمع المدني Media Monitoring Africa.

....

[495] https://padre.org.za

[496] https://www.real411.org

24. فريق العمل المعني بالمعلومات المُضلّلة في جمهورية كوريا (2018)

شكّل الحزب الديمقراطي الحاكم فريق عمل لتقديم الشكاوى على أساس التشهير والتحريض على العنف وغير ذلك من معلومات مضللة جدلية (Kajimoto, 2018; Sang-Hun, 2018).

25. وحدة التهديدات المختلطة في إسبانيا (2019)

أنشأت الحكومة الإسبانية وحدة «التهديدات المختلطة» التي تتكوّن من وزارات مختلفة في الفترة التي سبقت انتخابات الاتحاد الأوروبي والانتخابات العامة في عام 2019. وكان الهدف من الوحدة اتخاذ إجراءات استجابة مبكرة لضمان نزاهة الانتخابات، مع التركيز على الأمن السيبراني ورصد المعلومات المُضلّلة (ودحضها في بعض الأحيان) (Abellan, 2019).

26. التحقيق الخاص في تطوير هيئة الدفاع النفسي في السويد (2018)

في آب/أغسطس 2018، كلّفت الحكومة السويدية محققاً خاصاً بتحليل مقترحات لتطوير هيئة للدفاع النفسي وتقديمها. وكان الهدف من ذلك زيادة الوعي والقدرة على الصمود في مواجهة المعلومات المُضلّلة والتأثير العدائي (Swedish government, 2018). تماشياً مع ذلك، تصف الوكالة السويدية للطوارئ المدنية (MSB، المسؤولة عن الطوارئ وإدارة الأزمات) هدفها في مجال الدفاع النفسي بأنه «ضمان إرادة الشعب السويدي وقدرته على الصمود أمام الضغط والتّضليل من الخصم، في إطار مجتمع ديمقراطي منفتح، يتمتّع بحُريّة الرأي وحُريّة وسائل الإعلام -قدر الإمكان في ظل هذه الظروف القاسية- علماً أنه لا ينبغي فرض أي رقابة» (ترجمة أجراها المؤلفون)[497]. وهي تقدم نصائح لتعزيز الدراية الإعلامية بغرض تحديد كيفية نقد مصادر المعلومات.

27. مبادرة الدراية الإعلامية «تعلّم أن تُميّز» في أوكرانيا (2015)

أقامت وزارة التعليم الأوكرانية شراكةً مع المنظمة غير الحكومية الأمريكية IREX لوضع برنامج للدراية الإعلامية؛ هدفه تعليم المراهقين كيفية اكتشاف المعلومات المُضلّلة وخطاب الكراهية. وتلقّى البرنامج الدعم من سفارتَي المملكة المتحدة والولايات المتحدة في أوكرانيا (Ingber, 2019; Jankowicz, 2019).

وفي هذا السياق، اعتُمدت سياسة وزارة الإعلام الأوكرانية[498] في عام 2014 التي تقضي بـ«مُكافحة عُدوان المعلومات»، مما أدى، من بين أمور أخرى، إلى حجب بعض المواقع الأجنبية (انظر «إنفاذ القانون وتدخلات الدولة الأخرى»)، وإلى اقتراح قانونَين (على الرغم من سحبهما في وقت لاحق) من شأنهما تجريم التّضليل والسماح أيضاً بحظر وسائل إعلام (Jankowicz, 2019).

....

497 https://www.msb.se/psykologisktforsvar

498 http://mip.gov.ua/en/content/pro-ministerstvo.html

28. التحقيق الذي أجراه مجلس العموم البريطاني (لجنة الشؤون الرقمية والثقافة والإعلام والرياضة) بشأن المعلومات المُضلّلة و«الأخبار الزائفة» (2019)

أجرت لجنة التحقيق في الشؤون الرقمية والثقافة والإعلام والرياضة في مجلس العموم البريطاني تحقيقاً لمدة 18 شهراً طال فيس بوك وCambridge Analytica وAggregateIQ ونُشر في شباط/فبراير 2019. وتعكس التوصيات الواردة في التقرير تلك التي رفعتها اللجنة البرلمانية الدائمة الكندية وتقارير الحكومة الأسترالية، وتبرز خطر التّضليل على الديمقراطية وضرورة تعزيز قواعد الخصوصية والإعلان السياسي والقانون الانتخابي. وبالنسبة إلى تنظيم وسائل التواصل الاجتماعي على وجه التحديد، أوصى التقرير بضرورة «وضع مدونة أخلاقيات مُلزمة، تُشرف عليها هيئة ناظمة مستقلة، تحدّد ماهية المُحتوى الضار»، معتبراً أنّ هناك «مسؤولية قانونية واضحة لشركات التكنولوجيا للعمل ضد المُحتوى الضار وغير القانوني المتفق عليه على منصتها»، مع إلزام شركات التكنولوجيا بدفع ضريبة على الخدمات الرقمية بنسبة 2٪ (UK Commons Select Committee on Digital, Culture, Media and Sport, 2019, pp. 89-90). إلى ذلك، شدّد التقرير على أهمية الإلمام الرقمي، ووصفه بأنه «الركن الرابع للتعليم، إلى جانب القراءة والكتابة والرياضيات» (UK Commons Select Committee on Digital, Culture, Media and Sport, 2019, p.96).

في موازاة ذلك، أطلق مكتب مفوّض المعلومات في عام 2017 تحقيقاً حول إساءة استخدام البيانات الشخصية في الحملات السياسية. وقد تم تغريم عدد من المنظمات، بما فيها فيس بوك الذي دفع الغرامة القصوى التي بلغت 500,000 جنيه استرليني لانتهاكات طالت حماية البيانات (UK ICO, 2018). وأخيراً، قامت وزارة المملكة المتحدة للشؤون الرقمية والثقافة والإعلام والرياضة ووزارة الداخلية بوضع اللمسات الأخيرة على ورقة بيضاء حول الأذى على الإنترنت في حزيران/يونيو 2019، اقترحت من بين أمور أخرى على منصات التواصل الاجتماعي «واجب الرعاية» (انظر «الاقتراحات التشريعية»).

29. تحقيق لجنة الشؤون الخارجية في مجلس العموم البريطاني حول حُريّة الإعلام العالمية (الموضوع الفرعي هو المعلومات المُضلّلة) (2019)

ربط هذا التحقيق الذي أجرته لجنة الشؤون الخارجية في مجلس العموم البريطاني المعلومات المُضلّلة بالتهديدات التي تحدق بحُريّة الإعلام وسلامة الصحفيين. وناقشت اللجنة مشكلة «المعلومات المُضلّلة»، وكيف تهدد حُريّة الإعلام، والدور الذي يمكن أن تؤديه وسائل الإعلام الحرة رداً على ذلك»[499]. ونُشر التقرير النهائي للتحقيق في أيلول/سبتمبر (UKHouse of Commons Foreign Affairs Committee, 2019). وقد أقرّ بأن «وسائل الإعلام الحرة يمكن أن تكون أيضاً «الترياق» للتهديد المتزايد المتمثل في «التّضليل» –أي العرض المتعمد للأكاذيب كأخبار وقائعية من أجل تحقيق ربح شخصي أو سياسي أو تجاري– في حين أن وسائل إعلام غير حرة تواجه خطر تحوّلها إلى لسان حال التّضليل». وحذّر التقرير أيضاً من «... أنّ القوانين الواسعة أو المتسلّطة هي واحدة من أقوى الوسائل لإسكات الصحفيين، مع العلم أنّ القوانين المتعلقة بالتشهير أو الأمن القومي أو «نزع شرعية» «الأخبار الزائفة» أو المعلومات المُضلّلة أو الشائعات باتت بشكل متزايد معرّضة لسوء الاستخدام».

499 https://www.parliament.uk/business/committees/committees-a-z/commons-select/foreign-affairscommittee/news-parliament-2017/global-media-freedom-evidence-17-19/

30. تحقيق أجرته لجنة التحقيق في المخابرات في مجلس الشيوخ الأميركي في «حملات التدابير الناشطة والتدخل الروسي في الانتخابات الأميركية لعام 2016» (2017-2019)

أجرت لجنة التحقيق في المخابرات في مجلس الشيوخ الأمريكي تحقيقاً تناول مدى الانخراط الأجنبي في انتخابات عام 2016، ونشرت تقريرها في عدة مجلدات (U.S. Senate Select Committee on Intelligence, (2019a) Volume 1, July 2019, (2019b) Volume 2, October 2019 and (2019c) Volume 3, February 2020). ووجدت أدلة على التدخل من قبل شركة Internet Research Agency بشكل أساسي. ورفعت اللجنة توصيات إلى منصات التواصل الاجتماعي لتسهيل تبادل المعلومات مع الحكومة وسُلطات إنفاذ القانون والكونغرس بشأن التشريعات المتعلقة بشفافية الإعلان السياسي عبر الإنترنت. وحثّت السلطة التنفيذية على «إنشاء فريق عمل مشترك بين الوكالات لمتابعة ورصد استخدام الدول الأجنبية لمنصات التواصل الاجتماعي للتدخّل الديمقراطي ووضع إطار للردع» (U.S. Senate Select Committee on Intelligence, 2019d).

الاقتراحات التشريعية

تهدف غالبية الاقتراحات التشريعية الأخيرة (8 من أصل 13 تم تحليلها) إلى معالجة موضوع المعلومات المُضلّلة من خلال تنظيم المُحتوى، انطلاقاً من منظور مسؤولية المنصات الإلكترونية عن المعلومات المغلوطة أو مُحتوى الكراهية. وعلى غرار التحقيقات وفرق العمل، تركّز الاقتراحات التشريعية في بعض الأحيان على العملية الانتخابية. وتقوم اقتراحات تشريعية أخرى بتجريم نشر المعلومات المُضلّلة. ويمكن أن يؤدي ذلك إلى خطر استخدامها ضد الصحفيين المنشقّين، كما أشار إليه في عدة مناسبات ناشطون في مجال الحقوق الأساسية.

31. اقتراح قانون لإنشاء لجنة للتحقق من الأخبار الزائفة في الأرجنتين (2018)

وقت كتابة هذه الدراسة، كان هناك اقتراح قانون يهدف إلى إنشاء لجنة للتحقق من الأخبار الزائفة داخل الغرفة الانتخابية الوطنية يشقّ طريقه عبر البرلمان الأرجنتيني. فعند تلقي شكوى معينة، ستتحقق اللجنة من المعلومات المُضلّلة خلال الانتخابات الوطنية. والمعلومات المُضلّلة تُعرّف على أنها «تعليقات تبدو كأخبار تُبثّ على الإنترنت أو عبر وسائل أخرى، ويتم عادةً استحداثها للتأثير على الرأي السياسي أو كمزحة من دون أن تمتّ بصلة إلى الوقائع»، على الرغم من أن «أعمدة الرأي المُعبّر عنها بطرق تستند إلى المواقف الآيديولوجية أو المنطق المعبر عنه بشكل قاطع» معفاة من التدقيق (ترجمة غير رسمية، Argentina Bill on creating a Commission for the Verification of Fake News, 2018b).

وفي حال وجود معلومات مضللة، تقوم اللجنة بإخطار الغرفة الانتخابية الوطنية، التي تستطيع أن تطلب من منصات التواصل الاجتماعي وسم المنشور بـ«إشعار بالمصداقية المشكوك فيها» (ترجمة غير رسمية)، والتقليل من انتشارها، وإنشاء قاعدة بيانات للمعلومات المُضلّلة تكون متاحة مجاناً للجمهور. وتشمل العقوبات المقترحة؛ التحذيرات والغرامات وسحب الأهلية كمزوّد خدمات حكومي لمدة تصل إلى عشر سنوات، وفقدان الاستحقاقات أو النظم الضريبية الخاصة، وتعليق المنصة التي فشلت في نزع صفة الأولوية عن المُحتوى بعد صدور أمر بذلك لمدة عامين (Argentina Bill on creating a Commission for the Verification of Fake News, 2018b).

32. اقتراح لإنهاء ولاية سياسيين منتخبين بسبب التّضليل في تشيلي (2019)

في كانون الثاني/يناير 2019، اقترح خمسة من أعضاء مجلس الشيوخ التشيلي؛ قانوناً لإنهاء ولاية السياسيين المنتخبين إذا ثبُت أنهم قاموا بنشر معلومات مضللة عن المعارضين خلال الحملات الانتخابية أو قاموا بالترويج لها أو بتمويلها. ويهدف مشروع القانون إلى الحد من استخدام التّضليل لتحقيق مكاسب سياسية. وقد أُرسل الاقتراح إلى لجنة الدستور والتشريعات والعدالة والتنظيم في مجلس الشيوخ، ولكنه لم يُحلّل بعد (Chile Senate, 2019).

33. اقتراح قانون لمُكافحة خطاب الكراهية على الإنترنت في فرنسا (2019)

في عام 2019، بدأ البرلمان الفرنسي مناقشة قانون مُقترح لمُكافحة «التحريض على الكراهية والإهانات على أساس العِرق أو الدين أو الانتماء الإثني أو الجنس أو الميل الجنسي أو الإعاقة» (ترجمة أجراها المؤلفون) عبر الإنترنت. ينصّ اقتراح القانون، كما قدمه عضو البرلمان في آذار/مارس 2019، على ضرورة قيام المنصات عبر الإنترنت ذات الروابط العديدة على الأراضي الفرنسية بإزالة مُحتوى غير قانوني بشكل واضح في غضون 24 ساعة من تلقي شكاوى يرفعها مستخدمون أو بمنع الوصول إليه. وقد يُفضي عدم امتثال المنصات عبر الإنترنت إلى فرض غرامات. وبموجب القانون المقترح، سيتعيّن على المنصات الإلكترونية أيضاً توفير آليات لوسم المُحتوى والطعن في القرارات، فضلاً عن رفع تقارير شفافية تتناول إجراءاتها. ويكلّف الاقتراح هيئة الإذاعة الفرنسية رصد سلوك المنصات. وسوف يستند التشريع إلى قانون حُريّة الصحافة لعام 1881 (French Parliament, 2019; EurActiv, 2019).

34. تحديث القانون الألماني حول شبكات التواصل (2020)

في شباط/فبراير 2020، وافقت الحكومة الألمانية على حزمة تنظيمية من شأنها أن تكمل القانون المتعلّق بشبكات التواصل لعام 2017 (انظر «التشريعات المعتمدة»)، وهي تتطلّب قيام المنصات بالإبلاغ عن خطاب الكراهية غير القانوني للشرطة، وتوفير عناوين بروتوكول الإنترنت الخاصة بالمستخدمين. وعند كتابة هذه الدراسة، كان لا يزال على البرلمان أن يوافق على مشروع القانون. وكان من المتوقع تقديم نص ثانٍ لمراجعة القانون الأصلي في منتصف عام 2020، مع التركيز على إدارة المنصات للشكاوى (German BMJV, 2020a; German BMJV, 2020b).

35. التعديلات المقترحة على المبادئ التوجيهية لوسائط تكنولوجيا المعلومات في الهند (2018)

أصدرت وزارة الإلكترونيات وتكنولوجيا المعلومات الهندية مُسوّدة المبادئ التوجيهية (المعدّلة) لوسائط تكنولوجيا المعلومات للإدلاء بتعليقات عامة عليها في كانون الأول/ديسمبر 2018. وبموجب القسم الثالث المتعلق بالعناية الواجبة، اقترح تعديل أول أن تقوم الوسائط، بناءً على أمر من المحكمة أو إخطار من الوكالة الحكومية المختصة، في غضون 24 ساعة بإزالة المُحتوى الذي يُعتبر غير قانوني أو بتعطيل الوصول إليه– «تماشياً مع المادة 19 (2) من دستور الهند، لمصلحة سيادة الهند وسلامتها، أو أمن الدولة، أو العلاقات الودية مع الدول الأجنبية، أو النظام العام، أو الآداب أو الأخلاق، أو في ما يتعلق بازدراء محكمة أو التشهير أو التحريض على ارتكاب جريمة». ودعا تعديل ثانٍ الوسائط إلى «نشر أدوات مُميكنة قائمة على التكنولوجيا أو آلياتٍ مناسبة، مع ضوابط ملائمة، لتحديد المعلومات أو المُحتوى غير

القانوني وإزالته أو منع وصول الجمهور إليه بشكل استباقي» (Indian MeitY, 2018). وبعد فترة تشاور، نوقش الاقتراح بين الوزارات في تموز/يوليو 2019 ومن المتوقّع صدور نص معدّل في أوائل عام 2020 (Agrawal, 2019).

36. اقتراح لتنظيم شفافية الإعلانات السياسية على الإنترنت في آيرلندا (2019)

في تشرين الثاني/نوفمبر 2019، قدمت الحكومة الآيرلندية اقتراحاً تشريعياً لتنظيم شفافية الإعلانات السياسية المدفوعة على الإنترنت خلال فترات الانتخابات، تماشياً مع توصية المجموعة الآيرلندية المشتركة بين الإدارات المعنية بأمن العملية الانتخابية والمعلومات المُضلّلة (انظر: «التحقيقات وفرق العمل والمبادئ التوجيهية»، Irish Department of the Taoiseach, 2019).

37. تعديلات مُقترحة على قانون الانتخابات و«قوانين فيس بوك» في إسرائيل (2018)

قدّم الكنيست (البرلمان) الإسرائيلي مشروعات قوانين متعددة لإدخال تعديلات على القانون الانتخابي في مجال شفافية الإعلان الانتخابي ومُكافحة الدعاية الأجنبية خلال الانتخابات ونظر فيها. وكان عدد من مشروعات القوانين المتعلقة بإزالة المُحتوى غير القانوني عبر الإنترنت قد رُفعت أيضاً إلى الكنيست (Levush, 2019) ولكنها لم تُعتمد قبل انتخابات نيسان/أبريل 2019.

قبل انتخابات نيسان/أبريل 2019، اجتمع رئيس لجنة الانتخابات المركزية الإسرائيلية وقُضاة المحكمة العليا مع مسؤولي شركة فيس بوك لطلب حظر إعلانات انتخابية مجهولة المصدر، وقد امتثلت الشركة لهذا الطلب (Staff, 2019; Milano, 2019). ويأتي هذا الإجراء في أعقاب إحدى التوصيات الرئيسية للجنة فحص قانون الانتخابات (أساليب الدعاية) من أجل «توسيع نطاق تطبيق الأحكام الموضوعية للقانون ليشمل الإنترنت والمنصات الاجتماعية» (Levush, 2019).

38. مشروع قانون الحماية من الزيف والتلاعب على الإنترنت في نيجيريا (2019)

إنّ مشروع القانون النيجيري للحماية من الزيف والتلاعب على الإنترنت يُشبه إلى حد كبير قانون سنغافورة للحماية من الأكاذيب والتلاعب عبر الإنترنت.

ويقترح مشروع القانون حظر نقل بيانات كاذبة عن الوقائع، وصنع روبوتات الويب أو تعديلها للإبلاغ عن بيانات كاذبة عن الوقائع، وتوفير خدمات نقل بيانات كاذبة عن الوقائع في نيجيريا. ويستهدف مشروع القانون على نطاق واسع البيانات الوقائعية الكاذبة التي تضر بأمن نيجيريا والصحة العامة والسلامة والهدوء والشؤون المالية والعلاقات الودية مع البلدان الأخرى؛ والتي تؤثر على نتائج الانتخابات؛ وتُحرّض على العداوة أو الكراهية أو سوء النية بين مختلف فئات الناس؛ وتقلّص ثقة الجمهور بالحكومة.
تشمل العقوبات المفروضة على نشر بيان وقائعي كاذب الغرامات والحبس لمدة تصل إلى ثلاث سنوات. وعندما يُطلب من الأفراد القيام بذلك، يتعيّن عليهم نشر إشعار تصحيح والتوقّف عن نشر البيان الكاذب. ويقترح مشروع القانون أيضاً منح إدارة إنفاذ القانون النيجيرية (عن طريق لجنة الاتصالات النيجيرية) إمكانية إصدار أمر إلى مقدمي خدمات الإنترنت (الذين يُغرّمون في حال عدم الامتثال) بحجب الوصول إلى موقع معيّن على شبكة الإنترنت. ويشمل القانون أيضاً إمكانية الطعن أمام إدارة إنفاذ القانون ثم المحاكم (Nigeria Protection from Internet Falsehood and Manipulation Bill, Nigerian Senate, 2019; Paquette, 2019).

39. مشروع قانون مُكافحة المُحتوى الكاذب في الفلبين (2019)

رُفع مشروع قانون لمُكافحة المُحتوى الكاذب إلى مجلس الشيوخ الفلبيني في تموز/يوليو 2019. وقد أُحيل إلى لجنة الإعلام ووسائل الإعلام للنظر فيه. يحظر مشروع القانون «نشر معلومات يعرف ناشرها أو يعتقد بشكل معقول أنها كاذبة أو مضللة للجمهور»، عبر استخدام حساب وهمي لنشر مثل هذا المُحتوى المغلوط، أو تقديم خدمة وتمويل ذلك النشاط، أو عدم الامتثال للأوامر الواردة في القسم الخامس من القانون. ويجيز القسم الخامس من القانون لمكتب الجرائم السيبرانية التابع لوزارة العدل إصدار أوامر بتصحيح المعلومات المنشورة أو إزالتها أو حذفها. وتشمل العقوبات الغرامات والسجن لمدة تصل إلى اثني عشر عاماً. ويسمح مشروع القانون بالطعن أمام مكتب أمين سرّ وزارة العدل (Philippines Senate, 2019 Bill no 9, Anti-False Content Bill).

40. القانون المُقترح في جمهورية كوريا (2018)

تنظر الجمعية الوطنية في جمهورية كوريا في إدخال تعديلات على القانون تتراوح بين زيادة مسؤولية وسائط الإنترنت عن إدارة المعلومات المُضلّلة وتوسيع نطاق قوانين التشهير لتجريم المعلومات المُضلّلة (Sang-Hun, 2018; Corcoran,Crowley & Davis, 2019).

41. تعديلات مقترحة على القانون الجنائي في سريلانكا (2019)

بعد الهجوم الإرهابي الذي وقع في نيسان/أبريل 2019 مستهدفاً الكنائس والفنادق في عيد الفصح، تم حظر وسائل التواصل الاجتماعي في الأيام التالية على الهجوم من أجل الحد من التحريض على العنف ضد المُسلمين (انظر: «إنفاذ القانون وتدخلات الدولة الأخرى»). وفي أيار/مايو 2019، وافق مجلس الوزراء في سريلانكا على تعديلات طالت القانون الجنائي وقانون الإجراءات الجنائية لحظر الأخبار الزائفة وخطاب الكراهية الذي «يضرّ بالانسجام بين الأمم وبالأمن القومي». وتشمل العقوبات الغرامات والسجن لمدة تصل إلى خمس سنوات (AFP, 2019a; Sri Lanka Brief, 2019).

42. ورقة بيضاء حول الأذى على الإنترنت في المملكة المتحدة (2019)

أجرت حكومة المملكة المتحدة استشارة تناولت «ورقة بيضاء حول الأذى على الإنترنت» في عام 2019 وتضمّنت جلسات استماع محدودة[500]. وفي شباط/فبراير 2020، قدمت الحكومة ردها. وعلى الرغم من أن هذا الرد ليس في حد ذاته اقتراحاً تنظيمياً، إلا أنه يُرسي أساساً واضحاً لمزيد من الإجراءات التنظيمية الحكومية. وهو يلحظ واجب رعاية قانوني تتحمّله منصات التواصل الاجتماعي لحماية المستخدمين من المُحتوى الإرهابي الضار وغير القانوني، ومن المُحتوى المسيء للأطفال. وبناءً على ردود الأفعال من الاستشارة، شدّدت الحكومة على ما يلي: «لضمان حماية حُريّة التّعبير، ستحدد اللوائح التنظيمية توقعات متباينة من الشركات في ما يتعلق بالمُحتوى والنشاط غير القانونيين مُقارنةً بالسُّلوك الذي لا يُعتبر غير قانوني غير أنه قادر على التسبّب بالأذى»، وهو سلوك قد يشمل التّضليل. ويلحظ رد الحكومة على الورقة البيضاء أن تقوم هيئة ناظمة مستقلة (لربما تكون Ofcom، أي الهيئة الناظمة للاتصالات في المملكة المتحدة) برصد سلوك المنصات[501].

---

500 https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-whitepaper

501 https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response

43. مشروع قانون الهيئة التشريعية في ولاية تينيسي الأمريكية لتسجيل محطة CNN وصحيفة واشنطن بوست كوكلاء الحزب الديمقراطي «للأخبار الزائفة» (2020)

في شباط/فبراير 2020، ناقشت الهيئة التشريعية في ولاية تينيسي (2020) قرار مجلس النواب المشترك رقم 779، الذي قدمه أحد أعضائها (الجمهوري) «للاعتراف بمحطة CNN وواشنطن بوست كمصدر للأخبار الزائفة وإدانتهما لتشويه سمعة مواطنينا». ولخّصت لجنة الاستعراض المالي التابعة للجمعية العامة في ولاية تينيسي (2020) مشروع القانون على النحو التالي: «تقرَّر الاعتراف بأنّ محطة CNN وواشنطن بوست هما مصدراً للأخبار الزائفة وجزء من الجناح الإعلامي للحزب الديمقراطي، وتعتزم كذلك إدانة وسيلتَي الإعلام لتشويه سمعة مواطنينا والإشارة إلى أنهم أتباع ذوي تفكير ضعيف بدلاً من أن يكونوا أشخاصاً يمارسون حقوقهم التي دفع ثمنها قدامى المُحاربين بدمائهم». وبعد المناقشة، صدرت توصية بإقراره من قبل اللجنة القضائية في المجلس بعد تصويت سُجّل فيه 13 صوتاً مقابل 5 في شباط/فبراير 2020. وفي وقت كتابة هذه الدراسة، كان مشروع القانون يخضع لاستعراض تشريعي إضافي.

التشريعات المُعتمدة

تم تحديد 28 بلداً سنّت تشريعات تتعلّق بالتّضليل. تختار الحكومات معالجة مسألة المعلومات المُضلّلة، إما عن طريق تحديث اللوائح التنظيمية القائمة أو سنّ تشريعات جديدة. ويختلف نطاق التشريعات المعمول بها من قوانين إعلامية وقوانين انتخابية إلى قوانين حول الأمن السيبراني وقوانين جنائية. وتقوم اللوائح التنظيمية إما باستهداف مُرتكبي ما تعتبره السُّلطات تضليلاً (من أفراد ووسائل إعلام)، أو بإلقاء المسؤولية على شركات ومنصات التواصل عبر الإنترنت وفرض موجبات تقضي بمراقبة المُحتوى وإزالته في نهاية المطاف. وفي القوانين التي يُعرّف فيها التّضليل تعريفاً واسعاً أو في الأحكام الواردة في القوانين الجنائية، يبرز خطر كبير يتمثل في استخدام المعلومات المُضلّلة كوسيلة للسيطرة على المواطنين والصحفيين، مما قد يطرح بالتالي خطر الرقابة؛ إن لم تتوفّر مراقبة مستقلة.

44. قانون تمويل الأحزاب السياسية في الأرجنتين (2019)

في أيار/مايو 2019، أي قبل تنظيم الانتخابات العامة في تشرين الأول/أكتوبر 2019، تم إقرار عدة تعديلات على قانون الانتخابات الأرجنتيني تتناول شفافية تمويل الأحزاب السياسية. وقد منحت التعديلات التشريعية، من بين أمور أخرى، غرفة الانتخابات الوطنية الأرجنتينية ولاية الاحتفاظ بسجل لحسابات التواصل الاجتماعي والمواقع الإلكترونية للمرشحين والأحزاب السياسية، فضلاً عن سجل لشركات المسح واستفتاءات الرأي. وأُلزم أيضاً المرشحون والأحزاب السياسية بالكشف عن تمويل حملاتهم الإلكترونية، وتم ضمان التمويل العام الإضافي للتقارير الصحفية (Argentina Political Party Financing Law, 2019).

تأتي هذه التعديلات في أعقاب اتفاق استثنائي تمّ التوصّل إليه في آب/أغسطس 2018 في الغرفة الانتخابية الوطنية، حيث كان الطموح المتمثّل بإنشاء سجل للمرشحين السياسيين يهدف إلى المساعدة في الكشف عن الحسابات الزائفة (Rodriguez-Ferrand, 2019).

45. قانون الأمن الرقمي في بنغلاديش (2018)

في تشرين الأول/أكتوبر 2018، أقرت بنغلاديش (2018) قانون الأمن الرقمي، ليحلّ محل أجزاء من قانون تكنولوجيا المعلومات والاتصالات لعام 2006 ويعزّز قانون الأسرار الرسمية لعام 1923.

يُجيز القانون عمليات التفتيش والاستيقاف والقبض على الأشخاص من دون أمر قضائي، والسجن لمدة تصل إلى أربعة عشر عاماً، والغرامات على الجرائم السيبرانية، مثل الإرهاب السيبراني، وانتحال الصفة، والقرصنة، والوصول غير المشروع إلى معلومات الدولة وتوزيعها.

ويتضمّن القانون أيضاً أحكاماً بشأن تنظيم الخطاب عبر الإنترنت، مثل نشر «دعاية أو حملة» مناهضة للحرب التي خاضتها بنغلاديش في العام 1971 من أجل الاستقلال عن باكستان، و«المعلومات– البيانات الهجومية أو الكاذبة أو المسبّبة للخوف»، والمعلومات التي «تُعرقل المشاعر أو القيم الدينية» والتشهير بالأفراد.

وقد استُخدم القانون لطلب إزالة مُحتوى وتنفيذ عمليات قبض على مواطنين بسبب نشر شائعات ومعلومات كاذبة على المنصات الإلكترونية (Quadir, 2018; Alam, 2018).

46. قانون الإعلام في بيلاروس (نُقّح في عام 2018)[502]

في حزيران/يونيو 2018، أقرّ البرلمان البيلاروسي تعديلات على قانون الإعلام الخاص به ليتمكّن من مقاضاة الأفراد وحجب وسائل التواصل الاجتماعي والمواقع الإلكترونية بسبب نشر معلومات مضللة عبر الإنترنت (RFE/RL, 2018a). ويشمل التشريع فرض حظر على وسائل الإعلام الأجنبية، ولوائح تنظيمية أكثر صرامةً لتسجيل وسائل الإعلام واعتماد الصحفيين، وتحميل الناشرين على الإنترنت مسؤولية متابعة ورصد التشهير والتّضليل ومنعهما (EFJ, 2018).

47. القانون الرقمي في بنين (2017)

يتناول الفصل الرابع من قانون بنين الرقمي، الذي تم اعتماده في عام 2017، المُحتوى المسيء وجرائم الصحافة عبر الإنترنت، ويشمل جرائم مثل التحريض على الكراهية والعنصرية والعنف والإرهاب وإنكار الجرائم ضد الإنسانية والتشهير. وتتناول المادة 550 من القانون، على وجه التحديد، المضايقات عبر الإنترنت وتحظر نشر معلومات مضللة عن شخص ما. وتشمل العقوبات المفروضة على نشر المعلومات المُضلّلة الغرامات والسجن لمدة تصل إلى ستة أشهر (Benin Digital Code, 2017).

48. قانون تجريم التّضليل الانتخابي في البرازيل (2019)

في الفترة التي سبقت الانتخابات العامة للعام 2018، قُدّم عدد من القوانين الآيلة إلى تجريم التّضليل الانتخابي. وفي أيلول/سبتمبر 2019، أقرّ الكونغرس البرازيلي قانوناً يعدل قانون الانتخابات ويعرّف جريمة «الإبلاغ الافترائي لأغراض انتخابية»، ويفرض عقوبتي حبس و/أو سجن تتراوح مدّتها بين سنتين

[502] عندما تُستخدم عبارة «نُقّح في [عام كذا]»، فهذا يعني أن التعديلات التي أُدخلت على القانون ذي الصلة كانت تتعلّق بالمعلومات المُضلّلة على وجه التحديد.

وثماني سنوات. كما شكّل الكونغرس لجنة تحقيق برلمانية مشتركة للتحقيق في استخدام عملية تحديد السمات والتّضليل للتأثير على نتائج انتخابات عام 2018 (Redação Exame, 2019).

49. القانون الجنائي في بوركينا فاسو (نُقّح في عام 2019)

تجرّم التعديلات التي أُدخلت على القانون الجنائي في بوركينا فاسو في حزيران/يونيو 2019 «إضعاف معنويات» قوات الدفاع والأمن، وتقديم معلومات مغلوطة عن تدمير ممتلكات أو هجوم على أشخاص، ومعلومات تعرّض قوات الدفاع والأمن للخطر، ومعلومات تصف الإرهاب، وإهانة ذاكرة شخص متوفٍّ. وتشمل العقوبات الغرامات والسجن لمدة تصل إلى عشر سنوات. ويستطيع قاضٍ ما أن يأمر بمنع الوصول إلى مواقع إلكترونية أو مُحتوى يوزّع معلومات مغلوطة (RSF, 2019; Burkina Faso Constitutional Council, 2019).

50. توجيهات حول مُكافحة الأخبار الزائفة في كمبوديا (2018)

في الفترة التي سبقت انتخابات عام 2018، أصدرت الحكومة الكمبودية توجيهَين للسماح بحجب مواقع إلكترونية بسبب التهديدات الأمنية، وبتجريم نشر معلومات مضللة على الإنترنت. وتشمل العقوبات الغرامات والحبس لمدة تصل إلى عامين. وقد أُرغمت وسائل الإعلام (خارج الإنترنت، ولكن أيضاً إذا انتهت بـ kh.) على التسجيل في وزارة الإعلام الكمبودية (Lamb, 2018; Sovuthy, 2018). وهدّدت الوزارة بإلغاء التراخيص الإعلامية في حال نشر معلومات مضللة تُعرّض الأمن القومي للخطر. على ذلك، تمت إزالة المُحتوى المخالف أيضاً (Livsier, 2019).

51. القانون الجنائي (1967، منقّح) وقانون الأمن السيبراني والجرائم السيبرانية في الكاميرون (2010)

تحظر المادة 113 من القانون الجنائي الكاميروني «إرسال معلومات كاذبة يمكن أن تمسّ بالسُّلطات العامة أو الوحدة الوطنية أو نشرها» (Cameroon Penal code,1967, revised). وتشمل العقوبات الغرامات والحبس لمدة تصل إلى ثلاث سنوات. كما تم التصدي للتضليل من خلال أحكام أخرى، كتلك المتعلّقة بالتشهير والانفصال (CPJ, 2019a; Funke, 2019). وتُعاقب المادة 78 من قانون الأمن السيبراني والجرائم السيبرانية لعام 2010 نشر «أخبار دون إثبات صحتها أو تبرير وجود أسباب وجيهة للاعتقاد بصحتها» من خلال الوسائل الإلكترونية، مع فرض غرامات وعقوبة بالحبس لمدة تصل إلى سنتين. وتُضاعَف العقوبات إذا «ارتُكبت الجريمة بهدف تقويض السّلم العام» (ترجمة أجراها المؤلفون، Cameroon Cyber Security and Cyber Criminality Law, 2010).

52. قانون تحديث الانتخابات في كندا (2018)

في الفترة التي سبقت الانتخابات الفيدرالية لعام 2019، أقرّ البرلمان الكندي «قانون تحديث الانتخابات» (2018). يزيد القانون من شفافية الإعلانات الانتخابية للأحزاب السياسية والأطراف الثالثة، كما يفترض أن تقوم شركات التواصل الاجتماعي بتعزيز شفافية الإعلانات السياسية عبر الإنترنت. ونتيجةً لذلك، أجبرت كندا شركات التواصل الاجتماعي على «الاحتفاظ بسجل للإعلانات الحزبية والانتخابية التي نُشرت خلال الفترة التي سبقت الانتخابات وأثناء الانتخابات»[503].

[503] https://www.canada.ca/en/democratic-institutions/news/2019/01/encouraging-social-media-platforms-to-act.html

53. قوانين مُكافحة الشائعات في الصين (2016، 2017، 2018)

قامت الصين بتحديث عدة قوانين للتصدي للتضليل أو «الشائعات» وباعتمادها. فعلى سبيل المثال، يجرّم قانون الأمن السيبراني (2016) النشر المتعمّد للمعلومات الكاذبة والشائعات التي تقوّض النظام الاقتصادي والاجتماعي (Repnikova,2018). ويتناول القانون أيضاً على نطاق أوسع إدارة البيانات وأمن الشبكات. ويمكن الاطلاع على أحكام مماثلة في القانون الجنائي (المُحدّث في عام 2017) الذي يورد مسألة تخريب النظام العام والافتراء. ويمكن أن تشمل العقوبات السجن والاحتجاز الجنائي ومراقبة الحقوق السياسية والحرمان منها (People's Daily 2013; PKULaw, 2017;Repnikova, 2018).

علاوةً على ذلك، وفي ما يتعلق بتنظيم وسائط الإنترنت، تشير اللوائح التنظيمية الإدارية بشأن الخدمات الإخبارية على الإنترنت (2017) إلى أنه يتعيّن على المنصات الإلكترونية والخدمات الإخبارية إعادة نشر المصادر الرسمية المعتمدة من الحكومة وإنشاء روابط لها، في حين تُرغم اللوائح التنظيمية الإدارية بشأن معلومات المدونات الصغيرة (2018) المدونات الصغيرة (مثل: Weibo) على التحقُّق من هُويّة المدون واعتماد آليات «لمُكافحة الشائعات» تتيح رصد المعلومات المُضلّلة وحظرها ودحضها بشكل استباقي (Yang 2018; Repnikova, 2018).

54. القانون الجنائي (1981، منقّح)[504] وقانون الصحافة في ساحل العاج (نُقّح في عام 2017)

تحظر المادة 173 من القانون الجنائي في ساحل العاج (Cote d'Ivoire Penal Code 1981, revised, Article 173) «نشر أخبار كاذبة أو مواد ملفقة أو مزورة أو منسوبة بشكل مخادع إلى أطراف ثالثة، أو إشاعتها، أو الكشف عنها، أو إعادة إحيائها بأي وسيلة كانت» (ترجمة من قِبل المؤلفين). وتشمل العقوبات: الغرامات والحبس لمدة تصل إلى ثلاث سنوات. على ذلك، جرى توسيع نطاق المادة 97 من قانون النظام القانوني للصحافة (Cote d'Ivoire Penal Code 1981, revised. Article 97, 2017) لكي تشمل الصحافة الإلكترونية، وهي تحظر بالمثل «نشر مواد كاذبة أو مُلفّقة أو مُزوّرة أو منسوبة بشكل مخادع إلى أطراف ثالثة، أو إشاعتها، أو الكشف عنها، أو إعادة إنتاجها من خلال الصحافة»، علاوةً على فرض غرامات. (ترجمة أجراها المؤلفون) (Drissa, 2019).

55. قوانين مُكافحة الأخبار الزائفة في مصر (2018)

اعتمدت مصر قوانين متعددة تتعلق بحُريّة التّعبير والتصدّي للمعلومات المُضلّلة. أولاً، يسمح القانون رقم 180 للعام 2018 بشأن تنظيم الصحافة والإعلام للمجلس الأعلى لتنظيم الإعلام بتعليق وسائل إعلام ومواقع إلكترونية (تضم أكثر من 5000 متابع) تُشكّل تهديداً للأمن القومي، وتُعكّر السّلم العام، وتروّج للتمييز أو العنف أو العنصرية أو الكراهية أو التعصب وتنشر معلومات مضللة؛ وبمنعها وحجبها. ويفرض القانون أيضاً على وسائل الإعلام متطلبات إضافية من الناحية الإدارية ومن ناحية الترخيص (Sadek, 2019; TIMEP, 2019). ثانياً، ينصّ القانون رقم 175 لعام 2018 في شأن مُكافحة جرائم تقنية المعلومات على أنه باستطاعة سُلطة تحقيق حظر مواقع إلكترونية أو تعليقها إذا كانت تشكّل تهديداً للأمن القومي أو الاقتصاد الوطني، أو تحتوي على مُحتوى آخر يجرّمه قانون مُكافحة جرائم تقنية المعلومات. وتشمل

[504] عندما تُستخدم عبارة «[سنة]، منقّح»، فهذا يعني أن القانون ذا الصلة قد عُدّل (مرات متعددة في أغلب الأحيان) غير أن التعديلات الأخيرة لم تكن بالتحديد تتناول المعلومات المُضلّلة.

العقوبات أيضاً الغرامات والحبس لمدة تصل إلى سنتَين. ويتضمّن القانون أيضاً أحكاماً بشأن المسؤولية الجنائية لمقدمي الخدمات (Sadek, 2019;Balz & Mujally, 2019; Magdy, 2019). وأخيراً، تنصّ المادة 80 (د) من القانون الجنائي (1937، منقّح) على إمكانية فرض غرامات ومعاقبة «من أذاع عمداً في الخارج أخباراً أو بيانات أو إشاعات كاذبة حول الأوضاع الداخلية للبلاد؛ وكان من شأن ذلك إضعاف الثقة المالية في الدولة أو الإضرار بالمصالح القومية للبلاد» بالسجن لمدة تصل إلى خمس سنوات (Sadek, 2019).

56. قانون المعلومات غير الصحيحة في إثيوبيا (2020)

في شباط/فبراير 2020، أقرّ البرلمان الإثيوبي قانوناً لمُكافحة خطاب الكراهية ونشر المعلومات الكاذبة عبر الإنترنت. وقد يؤدي نشر مثل هذا المُحتوى على منصات لديها أكثر من 5000 متابع إلى غرامات أو عقوبة بالسجن لفترات مختلفة، تصل إلى خمس سنوات (Endeshaw, 2020).

57. قانون مُكافحة التلاعب بالمعلومات في فرنسا (2018)

في العام 2018، أصدر البرلمان الفرنسي قانوناً حول التلاعب بالمعلومات قبل فترة الانتخابات وخلالها («قانون مُكافحة الأخبار الزائفة»). ويستند التشريع إلى قانون حُريّة الصحافة لعام 1881. وقد طعن أعضاء من مجلس الشيوخ الفرنسي في القانون على أساس أنه يقيّد حُريّة التّعبير. ومع ذلك، صادق المجلس الدستوري على القانون وأصدره في كانون الأول/ديسمبر 2018 (Conseil Constitutionnel, 2018).

يعرّف القانون المعلومات المُضلّلة بأنها «اتهامات أو ادعاءات غير دقيقة أو مُضللة تهدف إلى تغيير صدق تصويت ما». وقبل ثلاثة أشهر من إجراء أي انتخابات، وبناءً على شكوى من السُّلطات العامة والمرشحين والجماعات السياسية والأفراد، يحقّ لقاضٍ مؤقت أن يتصرّف من خلال «اتخاذ تدابير مناسبة وضرورية» لوقف نشر معلومات مضللة. ويجب أن يصدر الحكم في غضون 48 ساعة من تلقي الشكوى. ويفرض القانون أيضاً «واجب التعاون» على المنصات الإلكترونية لتوفير آليات تتيح وسم المُحتوى، وضمان الشفافية الخوارزمية، وتعزيز المُحتوى الإخباري السائد، والشفافية الإعلانية، ومبادرات الدراية الإعلامية. ويمنح القانون الفرنسي أخيراً سلطة إضافية لهيئة الإذاعة الفرنسية لرصد سلوك المنصات وإلغاء تراخيص هيئات الإذاعة الأجنبية التي تنشر معلومات مضللة (France Fight against Manipulation of Information Law, 2018; Damiano Ricci, 2018).

58. القانون الألماني حول شبكات التواصل (2017)

اعتمد البرلمان الألماني قانون شبكات التواصل في حزيران/يونيو 2017. يلزم القانون منصات التواصل الاجتماعي التي تبغي الربح والتي تضمّ أكثر من مليوني مُستخدم مُسجّلين في ألمانيا؛ باتخاذ إجراءات ضد خطاب الكراهية والجرائم المنصوص عليها في القانون الجنائي الألماني (مثل: نشر الدعاية واستخدام رموز المنظمات غير الدستورية، والتحريض على العنف والجريمة والإرهاب، والإهانات الدينية، والتشهير، وتوزيع المواد الإباحية عن الأطفال وجلبها وحيازتها).

على ذلك، تُعتبر منصات التواصل الاجتماعي مُلزمة باتخاذ إجراءات واضحة لوسم المُحتوى ومعالجة الشكاوى، وإزالة مُحتوى «من الواضح أنه غير قانوني» أو منع الوصول إليه خلال 24 ساعة، ومُحتوى «غير قانوني» في غضون 7 أيام. وفي حال تلقّت أكثر من 100 شكوى في السنة التقويمية حول مُحتوى

غير قانوني، عليها نشر تقارير كل ستة أشهر عن كيفية تعاملها مع المُحتوى الذي جرى وسمه. ويتم إخطار الشخص الذي قدم الشكوى، إضافةً إلى المستخدم المتضرر، وتُرفع إليهما أسباب القرار (German NetzDG, 2017; German NetzDG English translation 2017).

59. قانون المعلومات والمعاملات الإلكترونية في إندونيسيا (نُقّح في عام 2016)

في تشرين الأول/أكتوبر 2016، نقّحت إندونيسيا قانون المعلومات والمعاملات الإلكترونية لعام 2008. يسعى القانون إلى تنظيم المُحتوى عبر الإنترنت ويأذن للحكومة بأن تأمر مزودي الأنظمة الإلكترونية بمنع الوصول إلى المُحتوى الذي يشكّل انتهاكاً للقانون، مثل المواد الإباحية أو الإرهاب أو الإهانة أو التشهير أو خطاب الكراهية، ولكن أيضاً أي مُحتوى آخر يُعتبر «سلبياً». وتشمل العقوبات الغرامات والسجن لمدة تصل إلى ست سنوات. ويشمل القانون أيضاً حق النسيان (أي حذف البيانات بناءً على حكم تصدره محكمة) (Molina, 2016;The Jakarta Post, 2016).

60. القانون الجنائي في كازاخستان (2014)

تحظر المادة 274 من القانون الجنائي في كازاخستان «نشر معلومات كاذبة عن علم، مما يستحدث خطر انتهاك النظام العام أو إلحاق ضرر جسيم بحقوق المواطنين أو المنظمات أو مصالحهم القانونية، أو مصالح المجتمع أو الدولة التي يحميها القانون» (ترجمة غير رسمية، Kazakhstan Penal Code, 2014). وتُفرض مجموعة من العقوبات، بما في ذلك الغرامات والأعمال الإصلاحية وما يصل إلى سبع سنوات من القيود المفروضة على الحُريّة، أو السجن (لمدة تصل إلى عشر سنوات إذا ارتُكبت الجريمة من قِبل جماعة إجرامية أو حصلت في أوقات الحرب أو حالات الطوارئ). ويمكن أن تؤدي المادة 130 من القانون الجنائي المتعلقة بالتشهير إلى فرض غرامات وأعمال إصلاحية وقيود على الحُريّة أو السجن لمدة تصل إلى ثلاث سنوات.

61. قانون حول سوء استعمال الحاسوب والجرائم الإلكترونية في كينيا (2018)

صدر القانون المتعلّق بسوء استعمال الحاسوب والجرائم الإلكترونية في كينيا في أيار/مايو 2018، وهو يشمل مجموعة واسعة من الجرائم المتّصلة بالحاسوب، بدءً بالوصول غير المصرّح به، والاعتراض والتدخّل، والتجسّس السيبراني والإرهاب السيبراني وصولاً إلى المواد الإباحية للأطفال، والمضايقات عبر الإنترنت، والاحتيال والتصيُّد وسرقة الهُويّة عبر الحاسوب.

يتعلق القسمان 22 و23 من القانون بالمنشورات الكاذبة، وهما يستهدفان «أي شخص ينشر عمداً بيانات كاذبة أو مضللة أو وهمية أو معلومات مغلوطة بهدف اعتبار البيانات سليمة والتصرّف على هذا الأساس، مع تحقيق ربح مالي أو بدونه». ويمكن أن تشمل العقوبات المفروضة على نشر هذه المعلومات المُضلّلة غرامة وحبس لمدة تصل إلى سنتين. وإذا كان نشر المعلومات الكاذبة «مخططاً له أو أدى إلى ذُعر أو فوضى أو عنف بين مواطني الجمهورية، أو من المرجح أن يفضي إلى تشويه سمعة شخص ما»، فتصل عقوبة السجن إلى عشر سنوات (Kenya Computer Misuse and Cybercrimes Act, 2018).

وقد طُعن في دستورية القانون أمام المحكمة، ويُتوقّع صدور حكم بهذا الخصوص في أوائل عام 2020. وفي انتظار ذلك؛ تم تعليق تنفيذ 26 قسماً من القانون، بما في ذلك الأقسام المتعلقة بالمعلومات المغلوطة (Itimu, 2019).

62. قانون مُكافحة الأخبار الزائفة في ماليزيا (مُلغى) (2019)

أصدرت ماليزيا قانون مُكافحة الأخبار الزائفة في نيسان/أبريل 2018. وقد عرّف القانون «الأخبار الزائفة» على نطاق واسع («أي أخبار ومعلومات وبيانات وتقارير، تكون مغلوطة كلياً أو جزئياً، سواء في شكل خصائص أو سمات معينة أو صور أو تسجيلات صوتية أو أي شكل آخر قادر على اقتراح كلمات أو أفكار»)، ونصّ على فرض عقوبات جاءت على شكل غرامات أو سجن لمدة تصل إلى ست سنوات. وسعى البرلمان الماليزي إلى إلغاء القانون مرتين: فعرقل مجلس الشيوخ المحاولة الأولى في آب/أغسطس 2018، لكنها نجحت في كانون الأول/ديسمبر 2019 (Free Malaysia Today, 2019). وبعد تغيير حكومي، اعتُبر قانون مُكافحة الأخبار الزائفة غير ضروري لأنّ القوانين القائمة (القانون الجنائي، وقانون الفتنة، وقانون المطابع والمنشورات، وقانون الاتصالات والوسائط المتعددة) تعالج بالفعل المعلومات المُضلّلة. وبموجب هذه المجموعة من التشريعات، يمكن أن تُفرض على الأفراد غرامات وعقوبات بالسجن لمدة تصل إلى سبع سنوات بسبب مجموعة واسعة من الإجراءات المتعلقة بالتّضليل، كالخطاب الذي «يضرّ بالنظام العام أو الأخلاق العامة أو الأمن، أو الذي يرجح أن يثير الخوف في أوساط الرأي العام، أو الذي يمس بالمصلحة العامة أو المصلحة الوطنية أو يُحتمل أن يمسّ بها»، علاوةً على التشهير والفتنة وخطاب الكراهية والتحريض على العنف (Buchanan, 2019).

63. قانون الاتصالات السلكية واللاسلكية (2013) والقانون الجنائي (1861، منقّح) في ميانمار

يحظر القسم 66 (د) من قانون الاتصالات السلكية واللاسلكية «ابتزاز أي شخص أو إكراهه أو تقييده بصورة غير مشروعة أو التشهير به أو إزعاجه أو التسبّب في تأثير لا مبرر له أو تهديده باستخدام أي شبكة اتصالات». وتشمل العقوبات الغرامات والحبس لمدة تصل إلى ثلاث سنوات (Myanmar Telecommunications Law, 2013). وقد استُخدم القسم 505 (ب) من القانون الجنائي أيضاً لكبح ما تعتبره السُّلطات معلومات مضللة. ويجرم هذا القسم «أي بيان أو شائعة أو تقرير» يرجّح أن يسبّب خوفاً في صفوف الجمهور أو قسم منه؛ بحيث يمكن بموجبه حثّ أي شخص على ارتكاب جريمة ضد الدولة أو ضد السَّكينة العام». وتشمل العقوبات الغرامات والحبس لمدة تصل إلى سنتين (Myanmar Penal Code, 1861,revised).

64. قانون تعديل الانتخابات في نيوزيلندا (2019)

على الرغم من أن القانون لا يتصدّى بشكل مباشر للمعلومات المُضلّلة، غير أنه قد يتمتّع بتأثير في هذا المجال. والقانون هذا قُدّم في كانون الأول/ديسمبر 2019 وسط مخاوف بشأن التدخل الأجنبي (بما في ذلك حملات القرصنة والتّضليل) خلال الانتخابات النيوزيلندية. يحظر القانون التبرعات السياسية الأجنبية التي تتجاوز 50 دولاراً نيوزيلندياً، ويفرض الشفافية في الإعلانات الانتخابية على جميع الوسائط (Ainge Roy, 2019)[505].

65. قانون الجزاء العُماني (1974، منقّح)

يتضمّن قانون الجزاء العماني أحكاماً تتعلق بنشر «خبر ارتكاب جريمة لم تُرتكب فعلاً، وهو يعلم أنها لم تُرتكب» و«إشاعات من شأنها النَّيل من هيبة الدولة». وتشمل العقوبات الغرامات والحبس لمدة تصل إلى ثلاث سنوات (Al Busaidi, 2019; Kutty, 2018).

[505] https://www.parliament.nz/en/pb/bills-and-laws/bills-proposed-laws/document/BILL_93304/electoral-amendment-bill-no-2

66. قانون منع الجرائم الإلكترونية في باكستان (2016)

يجرّم قانون منع الجرائم الإلكترونية الباكستاني (2016) أعمالاً مثل تمجيد الإرهاب، والإرهاب السيبراني (الذي يشمل التهديدات للحكومة وخطاب الكراهية)، والجرائم ضد كرامة الشخص أو تواضعه (مثل: المعلومات الكاذبة التي تنتشر لتخويف فرد أو إيذائه). وتشمل العقوبات الغرامات والسجن. ويسمح القانون أيضاً بإزالة المعلومات أو حجبها. وتشير وزارة الإعلام والإذاعة إلى دور قانون عام 2016 في زيادة الوعي ودحض المعلومات المُضلّلة من خلال حسابها على تويتر @FakeNews_Buster [506].

67. القانون الجنائي في الفلبين (نُقح في عام 2017)

تمّ تعديل المادة 154 من القانون الجنائي الفلبيني في عام 2017 لحظر نشر أخبار كاذبة «قد تعرّض النظام العام للخطر، أو تلحق ضرراً بمصلحة الدولة أو بسمعتها». وتشمل العقوبات الغرامات والحبس لمدة تصل إلى ستة أشهر (Philippines Revised Penal Code, Act no 10951, 2017).

68. تعديلات على قانون الإعلام وقانون المخالفات الإدارية تتعلّق بالأخبار الزائفة في روسيا (2019)

في آذار/مارس 2019، اعتمد الاتحاد الروسي قانونَين بغرض تعديل القانون الاتحادي حول الإعلام وتكنولوجيا المعلومات وحماية المعلومات (قانون الإعلام) وقانون الانتهاكات الإدارية.

تستند التعديلات التي أُدخلت على قانون الإعلام إلى التغييرات التي أُجريت عام 2016 والتي استحدثت نظام مسؤولية بالنسبة إلى مجمعات الأخبار في الاتحاد الروسي (مع أكثر من مليون مستخدم يومياً)، والتي منحت الخدمة الاتحادية الروسية للإشراف على الاتصالات وتكنولوجيا المعلومات ووسائل الإعلام (Roskomnadzor) سلطات رصد وحجب. وتستطيع كل من هيئات الدولة والمحاكم طلب حجب المُحتوى من خلال مجمعات الأخبار (Grigoryan, 2019; Richter, 2019). وتحظر التعديلات التي أُدخلت عام 2019 على قانون الإعلام نشر «معلومات هامة اجتماعياً غير موثوق بها» عبر الإنترنت، من شأنها أن تُشكّل تهديداً للمواطنين أو الممتلكات أو النظام العام و/أو الأمن العام أو النقل العام أو الصناعة أو الاتصالات» (Richter, 2019). وتسمح هذه التعديلات أيضاً للهيئة الناظمة للإعلام في الاتحاد الروسي Roskomnadzor، بناءً على طلب من المدعي العام ونوابه، إصدار أمر بإزالة معلومات مضللة على الإنترنت، وإذا لزم الأمر بحجب المواقع المخالفة (Grigoryan, 2019; Richter, 2019). ويشمل القانون إمكانية الطعن أمام المحكمة (Richter, 2019).

تحدد التعديلات التي أُدخلت على قانون المخالفات الإدارية بدورها الغرامات على نشر معلومات مضللة (Grigoryan, 2019; Richter, 2019). وأخيراً، تجدر الإشارة إلى أن القانون الجنائي للاتحاد الروسي يعالج أيضاً موضوع التّضليل، ويحظر على سبيل المثال إنكار الجرائم النازية وتزوير التاريخ. وتشمل العقوبات في هذا الصدد الغرامات والخدمة المجتمعية الإلزامية والحبس و/أو السجن (Grigoryan, 2019; Richter, 2019).

506 https://twitter.com/FakeNews_Buster

69. قانون الحماية من الأكاذيب والتلاعب عبر الإنترنت في سنغافورة (2019)

أصدر برلمان سنغافورة قانون الحماية من الأكاذيب والتلاعب عبر الإنترنت في أيار/مايو 2019. ودخل القانون حيز التنفيذ في تشرين الأول/أكتوبر 2019، مع منح استثناءات مؤقتة لبعض وسائط الإنترنت للامتثال (See Kit, 2019).

تحدّد صياغة القانون نطاق استهداف تلك البيانات الوقائعية الزائفة التي يُنظر إليها على أنها تضرّ بأمن سنغافورة وبالصحة العامة والسلامة والهدوء والشؤون المالية والعلاقات الودية مع البلدان الأخرى، وتؤثر على نتائج انتخابات ما، وتحثّ على مشاعر العداوة أو الكراهية أو الضغينة بين مجموعات مختلفة من الناس، وتقلّص ثقة الجمهور بالحكومة.

ويحظر القانون، في هذا الصدد، نشر بيانات وقائعية كاذبة عبر الإنترنت، وإنشاء روبوتات الويب لنشر بيانات وقائعية كاذبة أو تعديلها، وتوفير خدمات لنشر بيانات كاذبة عن الوقائع في سنغافورة. وتشير الرسائل الواردة من سلطات سنغافورة لأغراض وضع هذه الدراسة إلى ما يلي: «ينصّ القانون على ضرورة إجراء تصحيح كأداة أساسية له، الأمر الذي يتطلّب وضع إشعار إلى جانب المنشور أو المقالة الأصلية على شبكة الإنترنت يفيد بأنها تحتوي على بيان كاذب بشأن الوقائع وتوجيه المشاهدين إلى موقع حكومي على شبكة الإنترنت للإطّلاع على وقائع القضية. ويُعتبر القانون بمثابة آلية تضمن «حق الرد»، من دون إزالة المنشور الأصلي. والقصد من وراء ذلك هو السماح للقرّاء باتخاذ قرار بأنفسهم حول حقيقة الأمر».

وتضيف المراسلات ما يلي: «إذا لم يتم الامتثال لتوجيهات التصحيح؛ يجوز لحكومة سنغافورة أن تأمر مقدمي خدمات الإنترنت بمنع وصول المستخدمين النهائيين في سنغافورة إلى موقع محدد على الإنترنت. ويمكن أن يفضي عدم امتثال مزودي خدمة الإنترنت إلى غرامة. وفي الحالات القصوى التي تنطوي على خطر إلحاق ضرر جسيم، يجوز لحكومة سنغافورة إصدار توجيهات إلى الأفراد أو وسائط الإنترنت لجعل الأكاذيب على الإنترنت غير متاحة للمشاهدين في سنغافورة». وتضيف ما يلي: «بالإضافة إلى تأمين إمكانية الوصول إلى إشعارات التصحيح وزيادة وضوحها، تشمل التدابير التي تُتّخذ بموجب القانون تعطيل الحسابات المُفتعلة التي تضخّم الأكاذيب؛ وتشويه مصداقية مصادر الأكاذيب عبر الإنترنت؛ وحجب الحوافز المالية عن مصادر الأكاذيب عبر الإنترنت».

ويمكن تطبيق العقوبات المفروضة على نشر الأكاذيب عبر الإنترنت على الأفراد والمنظمات على حد سواء، وهي تشمل الغرامات وعقوبة بالسجن قد تصل مدّتها إلى عشر سنوات. وبناءً على توجيهات حكومة سنغافورة، يتعيّن على الأفراد أو وسائط الإنترنت نشر إشعار تصحيح والتوقّف عن نشر المعلومات الكاذبة. ووفقاً للرسائل التي وجّهتها سلطات سنغافورة إلى أمانة اليونسكو: «إذا لم يتم الامتثال للتوجيهات الواردة في القانون، يمكن توجيه أمر بحظر الوصول إلى مقدمي خدمات الإنترنت ولكن ليس إلى وسائط الإنترنت. غير أنه قد يُطلب من وسائط الإنترنت حجب الوصول إلى المُحتوى على منصتهم، إذا لم يمتثل مالك صفحة أو حائزها على المنصة لمتطلبات «الموقع المُشار إليه على الإنترنت» (ولكن هذا لا ينطبق على عدم الامتثال لتوجيهات القانون)».

ويتضمّن القانون أيضاً إمكانية الطعن أمام المحكمة. (Singapore POFMA, 2019)

70. قانون جرائم الحاسوب (نُقّح في عام 2017) وقانون الأمن السيبراني في تايلاند (2019)

يتصدّى قانون جرائم الحاسوب التايلاندي لجرائم مثل الوصول غير القانوني إلى نظام حاسوب وإلحاق الضرر به، والاعتراض غير القانوني للمعلومات والكشف عنها، فضلاً عن التشهير عبر الإنترنت. وفي ما يختصّ بالتّضليل، يحظر القسم 14 (2) «إدخال بيانات حاسوبية خاطئة في نظام حاسوب، على نحو يرجّح أن يضرّ بالأمن القومي أو السلامة العامة أو الأمن الاقتصادي الوطني أو البنية التحتية للصالح العام للأمة، أو يتسبّب بالذعر في صفوف الجمهور» (Thailand Computer Crime Act, 2017; Chitranukroh, 2017).

تتيح تعديلات عام 2017 لسلطات إنفاذ القانون إمكانية الوصول إلى أجهزة الحاسوب (مع مذكرة قضائية)، وتخوّل لجنة فحص بيانات الحاسوب (التي تعيّنها وزارة الاقتصاد الرقمي والمجتمع) وقف نشر بيانات الحاسوب غير القانونية (Thailand Computer Crime Act, 2017;Chitranukroh, 2017). وتشمل العقوبات الغرامات والسجن لمدة تصل إلى خمس سنوات. ويوسّع قانون الأمن السيبراني لعام 2019 نطاق وصول الحكومة عبر الإنترنت ويتيح لها، بالإضافة إلى ذلك، «استدعاء الأفراد للاستجواب والدخول إلى ممتلكات خاصة دون أوامر من المحكمة في حال «التهديدات السيبرانية الخطيرة» الفعلية أو المتوقّعة» (Tanakasempipat, 2019a).

71. قانون الأمن السيبراني في فيتنام (2018)

اعتمدت فيتنام قانون الأمن السيبراني في عام 2018، وهو قانون يحظر طائفة واسعة من الأفعال، مثل: خرق القوانين القائمة المتعلقة بالأمن القومي والسلامة العامة والنظام العام، وبناء معارضة لدولة فيتنام، والتحريض على العنف، و«نشر معلومات مضللة أو فاسدة أو إجرامية»، والتجسُّس الإلكتروني، والإرهاب، والقرصنة. وفي ما يتعلق بالتّضليل، ينصّ القانون على عقوبات تُفرض في حال «تقديم معلومات كاذبة، أو التسبّب في الارتباك بين المواطنين، أو إلحاق الأذى بالأنشطة الاجتماعية والاقتصادية، أو التسبّب بصعوبات تُعيق عمل وكالات الدولة أو الأشخاص الذين يؤدون واجبات عامة، أو انتهاك الحقوق المشروعة والمصالح التي تخصّ وكالات ومنظمات وأفراد آخرين» (Bich Ngoc, 2019; Chung Seck & Son Dang, 2019).

علاوةً على ذلك، يفرض قانون الأمن السيبراني على وسائط الإنترنت شرط توطين البيانات، من أجل توفير معلومات عن المستخدم للتحقيقات. كما أن مسؤولو الوسائط ملزمون بإزالة مُحتوى وحسابات غير قانونية أو بمنع الوصول إليها بشكل استباقي، وفي غضون 24 ساعة من تلقي طلب من قوة متخصّصة مسؤولة عن حماية الأمن السيبراني تكون تابعة لوزارة الأمن العام أو السُّلطات المختصة التابعة لوزارة الإعلام والاتصالات (Bich Ngoc, 2019; Chung Seck &Son Dang, 2019).

إنفاذ القانون وتدخّلات الدولة الأُخرى

يسلّط هذا القسم الضوء على أمثلة على إنفاذ اللوائح التنظيمية أو القوانين القائمة التي تُستخدم لمعالجة ما يُعتبر تضليلاً. وتشمل هذه الإجراءات: الغرامات، والقبض على الأشخاص، وقطع الإنترنت، وإزالة مواقع. وقد نُفّذت هذه التدابير بحقّ أفراد (بمن فيهم صحفيون وناشطون)؛ ووكالات إخبارية؛

ووسائل إعلام حكومية أجنبية اعتُبر أنها تنشر معلومات مضللة؛ وشركات تواصل عبر الإنترنت اعتُبرت مسؤولة عن الانتشار الهائل للمعلومات المُضلّلة. كما طُبّقت عمليات قطع الإنترنت لمنع انتشار معلومات مضللة في الفترات الانتخابية، وهو ما يمكن اعتباره بمثابة تدابير حادة (مفرطة/غير شاملة) للحد من الوصول إلى المعلومات. ولابدّ من الإشارة هنا إلى أن هذا المُحتوى يُعرض من دون أن تصدر اليونسكو أو لجنة النطاق الواسع أيّ حُكم أو تعليق على قرار اتُّخذ بموجب التشريعات السيادية، أو على شمولية المصدر/المصادر المذكور/ة.

72. البحرين

أخبار تفيد عن قبض على أشخاص وغرامات وعقوبات بالسجن على أساس التشهير ونشر معلومات كاذبة تهدّد أمن الدولة (BBC, 2018a; Associated Press, 2019).

73. بنغلاديش

في الأيام التي سبقت الانتخابات العامة في 30 كانون الأول/ديسمبر 2018، قُطعت خدمات الإنترنت عبر الهاتف المحمول «لمنع الشائعات والدعاية المُحيطة بالتصويت»، حسبما ذكر متحدّث باسم اللجنة الناظمة للاتصالات في بنغلاديش (Al Jazeera, 2018; Paul, 2018).

74. بنين

تم تطبيق المادة 550 من القانون الرقمي (انظُر: «التشريعات المعتمدة») لمُقاضاة صحفيين وفرض غرامة عليهم وسجنهم لنشر معلومات مُضلّلة عبر الإنترنت (Houngbadji, 2020;Fitzgibbon, 2020).

75. كمبوديا

نُقل أنّ مواطناً كمبودياً واحداً على الأقلّ أُدين وسُجن لانتهاكه قانون مُكافحة الأخبار الزائفة (انظُر: «التشريعات المعتمدة»)، لا سيما لنشر معلومات مضللة تهدّد الأمن القومي (Kongkea, 2019a & Kongkea, 2019b).

76. الكاميرون

أُوْدِعَ صحفيون السجن بتهم نشر أخبار كاذبة تهدّد أمن الدولة، على أساس القانون الجنائي وقانون الجريمة السيبرانية في الكاميرون (انظر: «التشريعات المعتمدة»). وامتثل المتهمون أيضاً أمام محكمة عسكرية (CPJ, 2019a;Funke, 2019).

77. جمهورية الصين الشعبية

من خلال قوانين مُكافحة الشائعات (انظر: «التشريعات المعتمدة») وغيرها من التشريعات ذات الصلة، تقوم الحكومة، بالتعاون مع وسائط الإنترنت، بحظر المُحتوى بشكل فعّال، وبإيقاف الحسابات، وبمقاضاة الأفراد لنشر ما اعتُبر أنه معلومات مضللة وشائعات تقوّض النظام الاجتماعي والاقتصادي والنظام العام (Qiu & Woo, 2018; Repnikova, 2018).

78. ساحل العاج

في عام 2019، فُرضت غرامة على سياسي في ساحل العاج وحُكم عليه بالحبس لمدة سنة واحدة بسبب تغريدة «تنشر أخباراً زائفة تُحرّض على العنف» (AFP, 2019c; BBC, 2019a).

79. مصر

تم إنفاذ الكثير من اللوائح التنظيمية، بما في ذلك قانون عام 2018 الذي ينظّم الصحافة والإعلام وقوانين مُكافحة الجريمة السيبرانية (انظر: «التشريعات المعتمدة»)، بغرض احتجاز أفراد وفرض غرامة عليهم، وحجب مواقع إلكترونية تنشر معلومات مضللة اعتبرتها السُّلطات تهدّد الأمن القومي (BBC, 2018b; Magdy, 2019).

80. ألمانيا

في سياق القانون الألماني حول شبكات التواصل (انظُر: «التشريعات المعتمدة»)، تم تغريم فيس بوك بمبلغ مليوني يورو من قِبل مكتب العدل الاتحادي الألماني في عام 2019 لعدم الشفافية في تقاريره عن الشكاوى المُقدمة والإجراءات المُتخذة لمعالجة خطاب الكراهية وغيره من الجرائم الجنائية (Prager, 2019; Zeit, 2019).

81. الهند

قُطع الإنترنت بشكل منتظم في الهند، وأحياناً لفترات طويلة (Gettleman, Goel & Abi-Habib, 2019; Burgess, 2018). وتسمح قواعد التعليق المؤقت لخدمات الاتصالات، التي اعتُمدت في عام 2017، للسُّلطات بتنظيم «التعليق المؤقت لخدمات الاتصالات بسبب الطوارئ العامة أو السلامة العامة» (Indian Ministry of Communications, 2017).

82. إندونيسيا

بالتعاون مع شركات التواصل عبر الإنترنت، تم إدراج نحو «مليون موقع إلكتروني» (Board, 2019) في القائمة السوداء بموجب قانون المعلومات والمعاملات الإلكترونية (انظر: «التشريعات المعتمدة») والتشريعات الأخرى ذات الصلة. ونُفّذت عمليات استيقاف وقبض على الأشخاص بتُهم نشر معلومات تنتهك القوانين الإندونيسية (Tapsell, 2019). وتمّ الإبلاغ عن تقييد الوصول إلى وسائل التواصل الاجتماعي لمنع «الخدع hoaxes» من خلال فرض قيود على القدرة على تحميل مقاطع الفيديو أو الصُّور (Beo Da Costa, 2019).

83. كازاخستان

في عام 2017، أُدين «فوربس كازاخستان» Forbes Kazakhstan وموقع ratel.kz بتهمة التشهير في دعوى قضائية. وتم تغريم الوسيلتَين الإخباريتَين وأُمرتا بإزالة مُحتوى تشهيري وباستصدار إشعار بالتراجع. وامتثل «فوربس كازاخستان» لهذه التدابير. أما موقع Ratel.kz، فقد دفع الغرامة، لكنه طلب توضيحاً من المحكمة بشأن المُحتوى الذي ينبغي إزالته. وعندما لم يتلقّ أيّ رد، لم يزل المُحتوى؛ ولم يصدر تراجعاً. وفي عام 2018، أمرت المحكمة بإيقاف موقع ratel.kz لمدة عام واحد لانتهاك قواعد تسجيل أسماء النطاقات

واستخدامها وتوزيعها في كازاخستان. وفي الحالتين، فُتحت تحقيقات بحقّ الوسيلتَين الإخباريّتين مردّها «نشر معلومات كاذبة عن علم» (Human Rights Watch, 2018a; RFE/RL, 2018b; Keller, 2019).

84. لاتفيا

في عام 2016، أمر مركز المعلومات الشبكية في لاتفيا بإغلاق الموقع المحلي لقناة Sputnik الإخبارية الأجنبية التابعة للاتحاد الروسي التي تُعتبر «أداة دعائية»، بعد أن لفتت وزارة الخارجية الانتباه إلى التغطية التي تُجريها القناة لأوكرانيا والإنكار الروتيني لسلامة أراضي هذا البلد (Latvian Public Broadcasting,2016; EurActiv, 2016).

85. ماليزيا

قامت ماليزيا بإنفاذ تشريعات مختلفة ذات صلة (بما في ذلك قانون مُكافحة الأخبار الزائفة الذي بات في الوقت الحالي مُلغًى) من أجل مقاضاة أفراد بسبب نشر معلومات مضللة. فعلى سبيل المثال، في كانون الثاني/يناير 2020، احتجزت لجنة الاتصالات والوسائط المتعددة الماليزية (2020) أربعة أشخاص يُشتبه في قيامهم بنشر أخبار مغلوطة عن جائحة فيروس كورونا بموجب القسم 233 من قانون الاتصالات والوسائط المتعددة.

86. ميانمار

استُخدم قانون الاتصالات السلكية واللاسلكية والقانون الجنائي (انظر: «التشريعات المُعتمدة») وتشريعات أخرى ذات صلة للحد من المُحتوى الذي اعتبرته السُّلطات مُضلّلاً في ميانمار (Associated Press, 2018; Schulman, 2019).

87. الاتحاد الروسي

قامت الهيئة الناظمة للإعلام في الاتحاد الروسي (Roskomnadzor) بحجب مُحتوى ومواقع اعتُبر أنها تقلّل من احترام السُّلطات في الاتحاد الروسي (Zharov, 2019) من خلال إنفاذ قانون الإعلام وتكنولوجيا المعلومات وحماية المعلومات (انظر: «التشريعات المعتمدة») وقوانين أُخرى ذات صلة (BBC, 2019b; Richter, 2019).

88. سنغافورة

استناداً إلى قانون الحماية من الأكاذيب والتلاعب عبر الإنترنت (انظر: «التشريعات المعتمدة»)، أمرت حكومة سنغافورة الأفراد والمنظمات بنشر إشعارات تصحيح بجانب المُحتوى الذي يُعتبر كاذباً (Palma, Munshi & Reed, 2020). ويتيح هذا القانون الطعن أمام المحكمة؛ وفي كانون الثاني/يناير 2020، قدّم حزب المعارضة، وهو الحزب الديمقراطي السنغافوري، الطعن الأول (Singapore Democratic Party, 2020)، لكن المحكمة العليا رفضته (AFP, 2020).

89. سريلانكا

في نيسان/أبريل 2019، أدى هجوم إرهابي على كنائس وفنادق في عيد الفُصح إلى ارتكاب أعمال عنف ضد المسلمين في سريلانكا. وتم حجب مواقع التواصل الاجتماعي في الأيام التالية للهجوم بهدف الحد من التحريض على العنف ضد المسلمين (Ellis-Petersen, 2019). وذكرت صحيفة واشنطن بوست أن تدابير مماثلة اتُّخذت في سريلانكا بعد اندلاع أعمال عنف ضد المُسلمين في عام 2018 (Romm, 2018; Dwoskin & Timberg, 2019).

90. تايلاند

قامت الحكومة التايلاندية بإنفاذ قانون جرائم الحاسوب (انظر: «التشريعات المعتمدة») على نطاق واسع للقبض على الأفراد بسبب نشر «أخبار زائفة» عبر الإنترنت (AFP, 2019b; Bangkok Post, 2019).

91. أوكرانيا

في العامَين 2017 و2018، فرضت أوكرانيا قيوداً على مُحرّك البحث Yandex، وشبكتَي التواصل الاجتماعي VKontakte وOdnoklassnik، و192 موقعاً آخر في أوكرانيا، في إطار تدابير الأمن القومي والعقوبات الاقتصادية ضد الاتحاد الروسي (Oliphant, 2017; Jankowicz, 2019).

Photo credits: Shutterstock, igorstevanovic

#DISINFODEMIC
#THINKBEFORESHARING
#SPREADKNOWLEDGE







